REMARKS

STATUS OF THE CLAIMS

Claims 49 – 51 and 55 – 78 were pending in this application. Claims 55, 59 – 66, 68, 73 – 74, and 77 – 78 have been cancelled without prejudice. Claims 56 – 58, 67, and 69 – 71 have been amended. Following entry of the amendments claims 49 – 51, 56 – 58, 67, 69 – 72, 75, and 76 will be pending and at issue.

SUPPORT FOR AMENDMENTS TO THE CLAIMS

Claims 56 – 58, 67, and 69 – 71 have been amended to revise their dependency from canceled claims 55, 65, 66, and 69 to pending claim 49. Support for the revised dependency can be found throughout the specification as filed, e.g., page 11, line 29-page 12, line 3, page 12, lines 18-26, page 13, lines 1-25, page 31, lines 13-31, page 33, line 1 through page 36, line 4.

The amendments to the claims therefore add no new matter.

STATEMENT REGARDING CLAIM CANCELLATIONS

To further prosecution, Applicants haves cancelled without prejudice claims 55, 59 – 66, 68, 73 – 74, and 77 – 78, rendering the pending rejections moot as applied to those claims. Applicants reserve the right to file subsequent applications claiming the canceled subject matter. In addition, the claim cancellations should not be construed as abandonment or agreement with the Examiner's position in the Office Action.

REJECTIONS UNDER 35 U.S.C. § 112, SECOND PARAGRAPH

Claims 49-51 and 55-78 stood rejected under 35 U.S.C. § 112, second paragraph as allegedly indefinite for recitation of the phrase "determining the functional effect." The Examiner states that "although the specification recites several examples of "functional effects" the skilled artisan could not be sure whether or not he or she was practicing the claimed invention because of the presence of such an ambiguous term." Office Action at page 2. Applicants respectfully traverse the rejection.

> The examiner's focus during examination of claims for compliance with the requirement for definiteness of 35 U.S.C. 112, second paragraph, is whether the claim meets the threshold requirements of clarity and precision, not whether more

> suitable language or modes of expression are available. When the examiner is satisfied that patentable subject matter is disclosed, and it is apparent to the examiner that the claims are directed to such patentable subject matter, he or she should allow claims which define the patentable subject matter with a reasonable degree of particularity and distinctness. Some latitude in the manner of expression and the aptness of terms should be permitted even though the claim language is not as precise as the examiner might desire. Examiners are encouraged to suggest claim language to applicants to improve the clarity or precision of the language used, but should not reject claims or insist on their own preferences if other modes of expression selected by applicants satisfy the statutory requirement.
>
> The essential inquiry pertaining to this requirement is whether the claims set out and circumscribe a particular subject matter with a reasonable degree of clarity and particularity. Definiteness of claim language must be analyzed, not in a vacuum, but in light of:
>
> (A) The content of the particular application disclosure;
>
> (B) The teachings of the prior art; and
>
> (C) The claim interpretation that would be given by one possessing the ordinary level of skill in the pertinent art at the time the invention was made.

MPEP section 2173.02, emphasis added.

Content of the application disclosure

The claims currently pending and at issue are restricted to methods in which "functional effects" are determined using an assay that comprises "contacting [a] compound with a cell expressing a sweet taste receptor comprising a T1R3 polypeptide and a T1R2 polypeptide." *See* Claim 49. The Examiner acknowledges that the specification recites several examples of "functional effects." Office Action at p. 2. These include, *e.g.*, measuring changes in cAMP, IP3 or Ca2+ (*see, e.g.,* specification at p. 6, lines 4-5), stimulation of enzyme activity such as phospholipase C and adenylate cyclase activity, and may be effectively measured, *e.g.*, by recording ligand-induced changes in $[Ca^{2+}]_i$ using fluorescent Ca^{2+}-indicator dyes and fluorometric imaging (*see, e.g.,* specification at p. 12, lines 18-26), measuring changes in voltage, membrane potential and conductance changes; ion flux, GPCR phosphorylation or dephosphorylation, *etc.* (*see, e.g.,* specification at p. 13, lines 8-25).

Teachings of the prior art

Applicants respectfully point out that the sweet taste receptors used for determining a functional effect of a compound are members of the large and well-studied G-protein-coupled receptors (GPCRs). The prior art teaches many assays for GPCR function, in addition to those described in the specification, including receptor phosphorylation as a marker of desensitization (*see, e.g.,* McDonald, *et al.* (2001), "βArrestins: New roles in regulating heptahelical receptors' functions," Cellular Signalling **13**:683-689)[1], GDP/GTP exchange and subsequent hydrolysis (*id.*), use of a green fluorescent protein (GFP) biosensor for detecting G Protein-coupled receptor activation (*see* Barak, *et al.*(1997), "A β-Arrestin/Green fluorescent Protein Biosensor for Detecting G Protein-coupled Receptor Activation," J. Biol. Chem. **272**:27497-27500)[2], as well as other reliable assays of GPCR function.

Claim interpretation by ordinarily skilled artisan

Finally, Applicants respectfully submit that the term "functional effect" when read in light of the claim as a whole, would reasonably be interpreted by an ordinarily skilled artisan at the time the invention was made to cover only those effects that are correlated with activation (including the desensitization that follows activation) or inhibition of signal transduction by the sweet taste GPCR.

For these reasons, Applicants respectfully submit that the claims define the subject matter with the reasonable degree of particularity and distinctness required under § 112, second paragraph, and request that the rejection be withdrawn.

REJECTIONS UNDER 35 U.S.C. § 112, FIRST PARAGRAPH

Claims 49-51, and 55-78 were rejected under 35 U.S.C. § 112, first paragraph, as allegedly containing subject matter which was not described in the specification in such a way as to enable one skilled in the art to which it pertains, or with which it is most nearly connected, to make and/or use the invention.

---

[1] A copy of McDonald, *et al.* (2001) is included as Appendix I.
[2] A copy of Barak, *et al.* (1997) is included as Appendix II.

Without agreeing with the Examiner's rejection but to expedite prosecution of this application, Applicants haves cancelled claims 55, 59-66, 68, 73-74, and 77-78 to focus on the enablement rejection as applied to cell-based assays.

The Examiner's position is that the specification is enabling for methods of identifying activators and inhibitors of sweet taste signal transduction, comprising a taste cell receptor composed of a heterodimer of SEQ ID NO: 9 and 15, wherein the receptor is present on the surface of a cell, and wherein the receptor is coupled to a Gα15 or Gα16 protein, but does not reasonably provide enablement for methods employing artificially constructed variants of SEQ ID NO: 9 and 15. Applicants respectfully traverse the rejection.

Essentially, the rejection, if maintained, would limit the claimed invention to exemplified embodiments, notwithstanding the breadth of Applicants' disclosure, the state of the art, and the high skill level of the artisan to whom the disclosure is directed. For reasons detailed below, Applicants respectfully submit that the full scope of the claims as amended is enabled, and request withdrawal of the rejection.

Applicants' invention is based on their pioneering discovery that a heterodimer comprising two members of the well-studied GPCR superfamily, *i.e.*, a T1R2 polypeptide and a T1R3 polypeptide, comprises a sweet taste receptor. That discovery, as taught by the Applicants in the specification, can be used to identify in a cell-based system, activators and inhibitors of sweet taste receptors.

Enablement is analyzed from the vantage point of an ordinarily skilled artisan having the benefit of the specification. If the ordinarily skilled artisan is able to practice the full scope of the claims without undue experimentation, then the enablement requirement is met. *See* MPEP § 2164.01. Factors to consider in this analysis come from *In re Wands*, 858 F.2d 731, 737, 8 USPQ2d 1400, 1404 (Fed. Cir. 1988) and are listed in MPEP § 2164.01(a): These factors include, but are not limited to:

(A) The breadth of the claims;

(B) The nature of the invention;

(C) The state of the prior art;

(D) The level of one of ordinary skill;

(E) The level of predictability in the art;

(F) The amount of direction provided by the inventor;

(G) The existence of working examples; and

(H) The quantity of experimentation needed to make or use the invention based on the content of the disclosure.

The breadth of the claims

The claims include structural limitations that effectively limit their scope to cell-based assays in which the cells express a "sweet taste receptor comprising a T1R3 polypeptide and a T1R2 polypeptide, wherein the T1R3 polypeptide has a greater than 90% amino acid sequence identity to SEQ ID NO: 15, 20, 23, or 25; and wherein the T1R2 polypeptide has a greater than 90% amino acid sequence identity to SEQ ID NO: 7, 8, or 9." Claim 49. The claims also include a functional limitation requiring that "the sweet taste receptor specifically binds a sweet compound." *Id.*

The nature of the invention

The invention is based on the Applicants' discovery that sweet taste receptors comprise heterodimeric G-protein coupled receptors (GPCRs) comprising a T1R3 and a T1R2 polypeptide. This pioneering discovery is put to advantage in the claimed invention through assays to identify compounds that activate or inhibit sweet taste signal transduction in taste cells.

The state of the prior art

The prior art includes a robust and well-developed literature dating back to the 1970s, on G-protein-coupled receptors (*see, e.g.* Lefkowitz (2004), "Historical review: A brief history and personal retrospective of seven-transmembrane receptors," *TRENDS* in Pharmacological Sciences **25**:413-422)[3]. The prior art establishes that the molecular properties of these GPCRs are well understood:

> Of the several large families of receptors, by far the largest, most versatile and most ubiquitous is that of the seven-transmembrane (7TM) receptors [also

[3] Lefkowitz (2004) is attached as Appendix III.

> referred to as seven-membrane-spanning receptors and G-protein-coupled receptors (GPCRs]. There are ~1000 genes encoding such receptors in the human genome, and these receptors regulate virtually all known physiological processes in mammals. . . .
>
> Moreover, we know a great deal about the molecular properties of these receptors: how they signal, how their function is regulated, and how they adapt to changing physiological circumstances and chronic time frames. Most of this information has been discovered during the past 30 years.

Lefkowitz (2004) at 413.

The level of one of ordinary skill

Applicants respectfully submit that the level of ordinary skill is quite high, and is represented by a person holding a Ph.D. or equivalent degree in a life-sciences discipline such as molecular biology or biochemistry, who is familiar with receptor biology and signal transduction.

The level of predictability in the art

Given the biological aspect of the invention, there is a degree of unpredictability that arises from the imperfectly understood correlation between protein structure and function. That level of unpredictability necessitates some degree of experimentation to make and use functional artificially constructed variants of the T1R2 and T1R3 polypeptides, starting from the sequences disclosed in the specification. The Examiner has pointed to two references in support of his argument that it would require undue experimentation to make and use sweet taste receptors having sequences at least 90% amino acid sequence identity to SEQ ID NOS: 15, 20, 23 or 25 (T1R3) or greater than 90% amino acid sequence identity to SEQ ID NOS: 7, 8, or 9.

First, the Examiner cites Bowie et al., 1990, Science 247:1306-1310 for the proposition that certain positions in sequence are critical to protein structure/function relationship, such as various sites or regions directly involved in binding, activity, and in providing the correct three-dimensional spatial orientation of binding and active sites. The Examiner correctly concludes that these regions can tolerate only relatively conservative substitutions or no substitution. *See* Office Action at p. 4.

Second, the Examiner cites to Guo, *et al.* PNAS 101 (9205-9210)2004 for the proposition that a systematic study of the tolerance that natural proteins have to amino acid sequence change

shows that "on average a single amino acid replacement had a 34% chance of inactivating a protein." Using this 34% inactivation probability per replacement, the Examiner estimates that the probability of generating an artificial, functional sequence having 90% sequence identity to SEQ ID NO: 9 or 15 as "astronomically low." *See* Office Action at p. 6. Applicants respectfully point out that the substitutions Guo, *et al.* studied were random (*see, e.g.,* Guo, *et al.* at p. 9205, right column, third line from bottom) and therefore not probative as to the likelihood of a skilled artisan's success in generating functional mutations that are other than random, such as by making, *e.g.*, conservative amino acid substitutions (explicitly taught in the specification at, *e.g.*, page 18, line 22 through page 19, line 2) or by targeting regions that are readily identified to be more likely tolerant of sequence change based on simple analysis of the multiple examples of sequences for T1R2 and T1R3 provided in the specification, as discussed in greater detail, below.

The amount of direction provided by the inventor

The Examiner states that "the specification has not provided guidance as to what properties of the allelic variants or sequence variants of the protein corresponding to SEQ ID NO:9 or 15 might be desired nor any guidance as to which amino acid substitutions, deletions or insertions to make to achieve any desired property." *See* Office Action at pp. 3-4. Applicants respectfully disagree. The properties that might be desired are those consistent with the practice of the claimed invention. To be useful in the practice of the claimed invention, allelic variants or sequence variants, must be capable of forming a sweet taste receptor that "specifically binds a sweet compound" and that can be used to "identify[ing] a compound that activates or inhibits sweet signal transduction." *See* Claim 49.

The Examiner also indicates that the specification fails to teach one of skill which amino acid substitutions, deletions or insertions to make that will preserve the structure and function of a protein corresponding to SEQ ID NO: 9 or 15. *See* Office Action at p. 4. Again, Applicants respectfully disagree. First, the specification discloses specific amino acid sequence changes among three alleles:

> The sequences of T1R3 in Sac-taster and non-taster mouse strains (C57BL/6 and 129/Sv) were determined from the genomic clones. The sequence of the entire

> coding region of the gene of other mouse strains that are sweet sensitive: SWR, ST, C57L, FVB/N and sweet insensitive: DBA/1Lac, DBA/2, C3H, AKR, BALB/c was determined from amplified genomic DNA (Jackson Laboratory). For SWR mice, T1R3 was also sequenced from amplified taste-tissue cDNA. Amongst the 11 inbred strains, we found two taster alleles (taster1: C57BL/6, C57L and taster 2: SWR, ST FVB/N) and a single non-taster allele (DBA/1Lac, DBA/2, C3H, AKR, BALB/c, 129/Sv). Taster 1 and taster 2 alleles differ from each other in six amino acid positions (P61L, C261R, R371Q, S692L, I706T, G855E; one of this G855E, was missed by (Kitagawa et al., 2001; Max et al., 2001) likely due to its inclusion in the primers used in their amplifications reactions). Non-tasters differ from taster 1 allele in six residues (A55T, T60I, L61P, Q371R, T706I, E855G), and from taster 2 in 4 amino acid positions (A55T, T60I, R261C, L692S).

Specification as filed at 59:27-60:5, emphasis supplied.

The specification also includes disclosure of T1R2 amino acid sequences for rat, mouse and human (SEQ ID NOS: 7, 8, and 9), and T1R3 amino acid sequences for human (SEQ ID NO: 15), mouse (SEQ ID NOS: 20, and 23) and rat (SEQ ID NO: 25).

Finally, the specification teaches that "[p]referably, the T1R3-comprising sweet taste receptor will have a sequence as encoded by a sequence provided herein or a conservatively modified variant thereof." Specification at p. 31, lines 13-14. Based on these sequence disclosures, the suggestion to use a conservatively modified variant, and explicit disclosure of conservative amino acid substitutions, an ordinarily skilled artisan would readily be able to carry out a sequence alignment (as, *e.g.*, known to one of ordinary skill and, *e.g.*, taught by the specification at page 16, line 10-page 17, line 16) and identify residues tolerant to conservative substitution amongst the disclosed sequences, as well as residues that are absolutely invariant and so unlikely to tolerate any substitution. Such exemplary alignments have been carried out by Applicants and are included as Appendices IV (T1R2 sequences) and V (T1R3 sequences) to this paper. [4]

---

[4] The invariant residues and conserved substitutions are indicated with symbols below the aligned sequences found in Appendices IV and V. Residues that are identical in each aligned sequence are indicated with the symbol "*", residues that are conservatively substituted are indicated with the symbol ":", and residues that are semi-conservatively substituted are indicated with the symbol ".". *See* ClustalW help file (Appendix VI) at p. 4 under "Consensus Symbols" heading.

The alignments provide additional evidence that the ordinarily skilled artisan would be able to make functional T1R2 and T1R3 polypeptides sequences that are "at least 90% identical" to the exemplified sequences in view of the exemplified sequence divergence. *See* the last column in the "Scores Table" on the first page of Appendices IV, showing identities amongst exemplified sequences ranging from 69% to 99%. Thus, by exercise of no more than ordinary skill, an artisan having the benefit of the disclosure would be able to construct additional functional sequences within the scope of the claimed invention.

The Examiner also has argued that the claims are essentially single means claims, citing *In re Hyatt*, 708 F.2d 712, 218 USPQ 195 (Fed. Cir. 1983), *Fiers v. Sugano*, 984 F.2d 164, 25 USPQ2d 1601 (Fed. Cir. 1993), and MPEP § 2164.08(a). Applicants respectfully submit that these cases are easily distinguished from the instantly claimed invention. As for *In re Hyatt*, the instant claims do not fall under 112, paragraph 6, as they recite structure (i.e., an amino acid sequence that is at least 90% identical to a disclosed sequence), and they are recited as part of a combination (i.e., a receptor that comprises a combination of a T1R2 and a T1R3 polypeptide), along with method steps of contacting a compound with the receptor. As for *Fiers v. Sugano*, the key distinction is that Applicants were in possession of and disclosed actual sequence data (three examples of T1R2 sequences and five examples of T1R3 sequences) to support their claimed invention. Revel's specification disclosed no sequence data whatsoever. *See Fiers v. Sugano*, 984 F.2d at 1170-1171.

The existence of working examples

The specification includes working examples teaching molecular cloning of T1R3, generation of transgenic mice for behavioral assays, heterologous expression of T1Rs in HEK-293 cells, and functional testing of heterologously-expressed T1Rs using calcium imaging. The specification also provides actual working examples showing that two different promiscuous G proteins (Gα15 and Gα16) can be used to couple the sweet taste receptors to a functional readout ($[Ca^{2+}]_i$ changes) in heterologous cells. *See* specification at pp. 59-63.

Quantity of experimentation needed

As discussed above, the claimed invention is directed to the field of receptor biology. As such, there necessarily is a degree of unpredictability that necessitates some experimentation to practice the full scope of the claimed invention. Sequence analysis as between two different taster and one non-taster allele provided in the specification, guidance to preferable, conservative amino acid substitutions, and simple sequence alignments taught by the specification, guide the ordinarily skilled artisan in the choice of amino acid changes to make in constructing T1R2 and T1R3 sequences at least 90% identical to those disclosed, that are useful for practicing the claimed methods.

In summary, Applicants respectfully submit that analysis of the Wands factors leads to the favorable conclusion that the full scope of the pending claims is enabled, and request withdrawal of the rejection based on §112, first paragraph.

CONCLUSION

Withdrawal of the pending rejections and reconsideration of the claims are respectfully requested, and a notice of allowance is earnestly solicited. If the Examiner has any questions concerning this Response, the Examiner is invited to telephone Applicants' representative at (415) 875-2413.

Respectfully submitted,
CHARLES ZUKER

Dated: July 28, 2015

By: Michael J. Shuster

Michael J. Shuster, Ph.D.
Reg. No.: 41,310
Fenwick & West LLP
Silicon Valley Center
801 California Street
Mountain View, CA 94041
Tel.: (415) 875-2316
Fax.: (650) 938-5200

Appendix I

ELSEVIER

Cellular Signalling 13 (2001) 683–689

CELLULAR SIGNALLING

Review article

βArrestins: New roles in regulating heptahelical receptors' functions

Patricia H. McDonald, Robert J. Lefkowitz*

Departments of Medicine and Biochemistry, Howard Hughes Medical Institute, Duke University Medical Center, Box 3821, Durham, NC 27710, USA

Received 28 February 2001; accepted 2 April 2001

Abstract

The last few years have seen a marked expansion in appreciation of the diversity of roles played by the βArrestins in regulating GPCR functions. Originally discovered as molecules that desensitize such receptors, the roles of βArrestins have expanded to include acting as signalling adapters or intermediates that recruit other key molecules to the GPCRs in an agonist-regulated fashion. For example, interactions with components of the endocytic machinery, such as clathrin, the adapter protein AP-2 and the *N*-ethylmaleimide sensitive fusion protein (NSF), demonstrate the ability of βArrestins to act as adapters to facilitate the clathrin-mediated endocytosis of certain members of the GPCR family. βArrestins have also been shown to serve as signalling molecules. The Ras-dependent activation of ERK1/2 may involve the βArrestin-dependent recruitment of c-Src to the β2-adrenergic receptor (β2-AR). More recently, βArrestins have been shown to act as molecular scaffolds that coordinate the assembly of certain MAP kinase complexes that lead to the stimulation of either ERK1/2 or JNK3. Finally, long-term accumulation of arrestin–rhodopsin complexes, in photoreceptor cells has been shown to trigger apoptosis. © 2001 Elsevier Science Inc. All rights reserved.

Keywords: βArrestin; G-protein-coupled receptor; GPCR; Scaffold; Mitogen activated protein kinase; MAPK

1. Introduction

G-protein-coupled receptors (GPCRs) represent a large and functionally diverse family of receptors whose primary function is the transmission of information from the extracellular environment to the interior of the cell. They include receptors for many different signalling molecules such as hormones, neurotransmitters, chemokines, taste and odorant molecules and even photons and calcium ions. GPCRs share a common molecular structure comprising an extracellular N-terminal domain, seven transmembrane spanning domains and a cytoplasmic C-terminal domain linked by three extracellular and three intracellular polypeptide loops. Extracellular signals are transduced to the cytoplasm via the heptahelical receptor by interaction of the receptor with its cognate heterotrimeric guanine nucleotide binding (G) protein. Agonist binding to the receptor induces a conformational change in the receptor and promotes the exchange of GDP for GTP binding to the G-protein. This causes dissociation of the G-protein into its α and βγ subunits leading to activation of downstream effectors by both the Gα and Gβγ subunits [1].

Interfering with the receptor's ability to activate its G-protein is one mechanism whereby attenuation or termination of receptor signalling occurs, a process termed desensitization. Desensitization of select GPCRs such as β2-adrenergic receptors (β2-ARs) takes place during short-term exposure to agonist and involves the rapid phosphorylation of the activated receptor by a specific G-protein-coupled receptor kinase (GRK). GRK-mediated phosphorylation of the agonist-occupied receptors promotes binding of arrestin proteins, which when bound elicit uncoupling of the receptor from its G-protein, which effectively terminates the signalling [2–4].

The founder member of the arrestin family, visual arrestin, was originally identified in the context of 'adaptation' or desensitization of the photoreceptor rhodopsin [5]. The nonvisual arrestin, βArrestin1 (Arrestin 2) was subsequently identified following the observation that the ability of GRK2 preparations to desensitize the β2-AR diminished during purification of GRK2, suggesting the existence of a

* Corresponding author. Tel.: +1-919-684-2974; fax: +1-919-684-8875.

E-mail address: lefko001@receptor-biol.duke.edu (R.J. Lefkowitz).

0898-6568/01/$ – see front matter © 2001 Elsevier Science Inc. All rights reserved.
PII: S0898-6568(01)00203-0

cofactor that was lost during purification [6]. This protein was postulated to be an arrestin-like protein as visual arrestin enhanced the inactivating effects of GRK2 on β2-AR signalling. Another nonvisual arrestin, βArrestin2 (Arrestin 3) was later cloned from a rat brain cDNA library [7]. The Arrestin family now includes at least four members, several of which undergo alternative splicing. Expression of some forms (visual arrestin and cone arrestin) is confined to the retina where they function specifically in the inactivation of rhodopsin. Other forms, βArrestin1 and βArrestin2, are ubiquitously distributed and appear to play a role in the uncoupling of many GPCRs [2–4].

βArrestins were once believed to be exclusively involved in receptor desensitization. However, over the last few years, a variety of other roles of βArrestins in heptahelical receptor regulation have been uncovered. In addition to their role in inactivating receptors, βArrestins have also been shown to be involved in receptor internalization [8]. In this process, βArrestins bind to components of the endocytic machinery such as clathrin [9], AP-2 [10] and the *N*-ethylmaleimide sensitive fusion protein (NSF) [11] and act as adapters to facilitate the clathrin-mediated endocytosis of certain members of the GPCR family such as the β2-ARs. GRK phosphorylation of the receptor and subsequent βArrestin binding not only serves to desensitize receptors and lead to their internalization via clathrin-coated pits but also plays additional roles in signalling. This review summarizes earlier studies describing the role of βArrestins in desensitization and internalization of certain heptahelical receptors and highlights more recent findings indicating that by acting as molecular adapter/scaffold molecules, βArrestins can trigger additional signalling events.

2. Desensitization

Receptor desensitization is a process that terminates or dampens receptor signalling. Several molecular mechanisms of desensitization have been delineated. One common mechanism involves rapid phosphorylation of the activated receptor by a specific GRK followed by βArrestin binding [2–4]. The function of arrestins in desensitization has been extensively studied using purified proteins in reconstituted systems. Although visual arrestin modestly enhances GRK2-mediated desensitization of the β2-AR, βArrestin inhibits signalling of the GRK2-phosphorylated β2-AR by more than 80% [7].

A variety of strategies have been used to study the arrestins in more intact systems. For example, in intact cells, desensitization of the β2-AR [12] and β1-AR [13] to high concentrations of the β-AR agonist, isopreterenol, is significantly enhanced by overexpression of GRK2 and βArrestin. In addition, incubation of olfactory cilia with neutralizing antibodies to βArrestin2 and GRK2 attenuates desensitization of odorant receptors [14]. Direct evidence for the functional significance of arrestins in intact animals comes from studies in both *Drosophila* and mice. Mutations in *Drosophila* arrestin genes cause a defect in the ability of the animal to inactivate meta-rhodopsin, and 5 days of light/dark exposure results in retinal degeneration [15]. The importance of βArrestins in GPCR signalling is demonstrated by findings in both βArrestin1 and βArrestin2 knock-out mice. βArrestin1-deficient mice display increased cardiac contractility in response to β-adrenergic receptor agonists [16]. βArrestin2-deficient animals have no gross phenotypic abnormalities, however, they demonstrate a remarkable potentiation and prolongation of the analgesic effect of morphine compared to wild-type littermate controls, demonstrating that μ opioid receptor desensitization is impaired [17]. Furthermore, these animals do not become tolerant to morphine even after repeated administrations, although, like wild-type animals, they do become physically dependent [18].

3. Internalization

In addition to uncoupling of the receptor from its G-protein, this process may serve as a prelude to receptor downregulation [19] and also contribute to resensitization of receptors. Sequestered receptors are ultimately either dephosphorylated and recycled back to the plasma membrane (resensitization) or targeted for degradation (down regulation). Many heptahelical receptors such as the β2-AR utilize classical clathrin-coated vesicle pathways for internalization. As described earlier, phosphorylation of the receptor by a GRK and βArrestin binding are crucial steps in this process [2–4].

It has been demonstrated that mammalian βArrestin1 and 2 (but not visual arrestin) bind directly and stoichiometrically to clathrin, the major structural component of the clathrin-based endocytic machinery [20–22]. βArrestins appear to possess the essential functional properties of an adapter molecule by interacting with high affinity with both receptors and clathrin. βArrestin mutants that are defective in binding to either receptors or clathrin show a reduced ability to promote β2-AR internalization. Moreover, a mutated form of the β2-AR (Y326A), which is a poor substrate for GRKs, cannot be internalized. However, internalization can be restored by overexpressing either GRK2 [23] or βArrestin1 [8]. The essential role of βArrestins in the internalization process is further confirmed by the finding that 'dominant negative' forms of βArrestin1, namely V53D [8] and S412D [24] dramatically impair receptor sequestration.

The function of βArrestin1 in GPCR internalization is regulated by phosphorylation/dephosphorylation of the βArrestin1 molecule. Cytoplasmic βArrestin1 is constitutively phosphorylated on a carboxyl-terminal serine (Ser-412) [24] and is recruited to the plasma membrane upon agonist stimulation of the receptor, where it becomes rapidly dephosphorylated. Dephosphorylation of βArrestin1 is required for clathrin binding and the subsequent targeting

Table 1
βArrestin interacting proteins (non-GPCR)

Binding partner	Arrestin isoform	Reference
Clathrin	βArr2>βArr1>Arrestin	[8,17–19]
NSF	βArr1, βArr2	([9], unpublished data)
AP2	βArr1, βArr2	[10,22]
c-Src	βArr1	[27,29]
ASK1	βArr1, βArr2	(unpublished data, [24])
JNK3	βArr2	[24]
Raf1	βArr1, βArr2	[23,31]
ERK	βArr1, βArr2	[30,31]

of receptors to clathrin-coated pits, but not for receptor binding and thus receptor desensitization. It is interesting to note that other members of the arrestin family (visual arrestin, βArrestin2 and other splice variants) do not possess the Ser-412 residue.

The ability of βArrestins to promote GPCR internalization does not arise solely as a consequence of the ability to bind clathrin. Recent studies suggest that additional βArrestin-binding proteins are involved in this process. For example, other endocytic proteins such as the adapter molecule AP-2 [10] and NSF [11] have also been shown to play a role in the internalization of GPCRs (Table 1). The adapter protein AP-2 plays a critical role in the recruitment of clathrin and the assembly of clathrin lattices that constitute the coat of the internalized membrane during endocytosis. It has been demonstrated that upon agonist activation of the β2-AR, βArrestins translocate to receptors at the plasma membrane where the receptor–βArrestin complex colocalizes with clathrin and AP-2 in clathrin-coated pits. AP-2 has been shown to interact directly with βArrestin1 and 2. Moreover, βArrestin mutants lacking the AP-2 binding motif are also defective in their ability to target receptor–βArrestin complexes to clathrin-coated pits [25]. Interestingly, clathrin and AP-2 bind to distinct regions of βArrestin giving rise to the possibility that these two proteins can bind independently and simultaneously. The binding of both proteins have been shown to be critical events for the internalization process, however, the order of events still remains to be elucidated.

A yeast two-hybrid screen identified NSF (an ATPase whose catalytic activity has been demonstrated to be essential in intracellular trafficking) as a βArrestin1 binding partner. In vitro binding data demonstrates that βArrestin1 preferentially interacts with the ATP bound form of NSF and that overexpression of NSF significantly enhances β2-AR sequestration. In addition, overexpression of NSF rescues the dominant negative effect of a βArrestin1 phosphorylation mutant (S412D; a mutant that mimics a constitutively phosphorylated form of βArrestin1 [24]) by restoring normal sequestration of the β2-AR [11], thus, further implicating NSF in the internalization process. Therefore, in addition to the βArrestin–clathrin interaction, the βArrestin–AP-2 interaction and the βArrestin–NSF interaction appear to be critical to the β2-AR internalization process. Fig. 1 outlines the role of βArrestin in the desens-

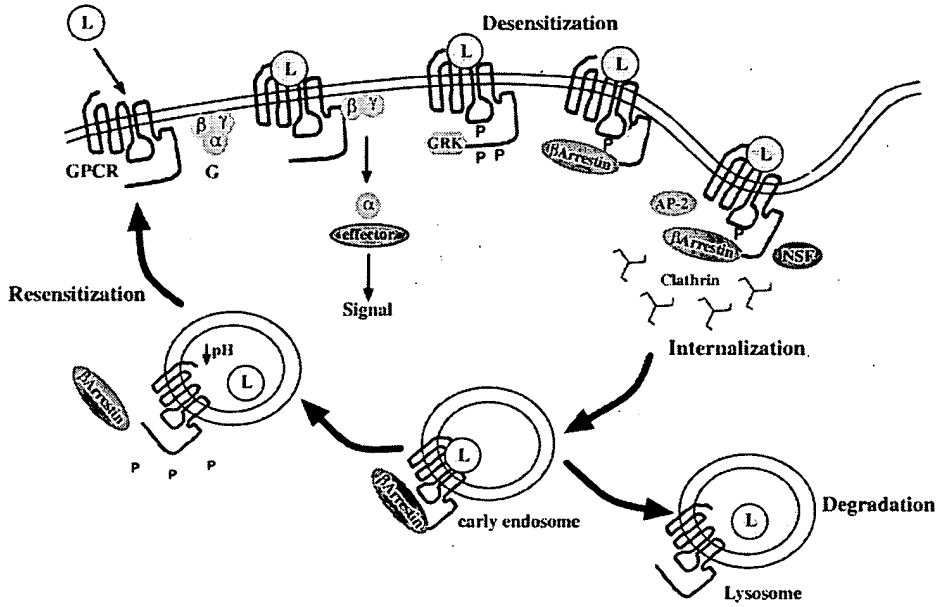


Fig. 1. βArrestin-dependent trafficking of heptahelical receptors. Ligand (L) or agonist binding to receptor is followed by phosphorylation of the receptor by a G-protein-coupled receptor kinase (GRK), which facilitates βArrestin binding and uncoupling from G-proteins leading to receptor desensitization. The ligand–receptor complex is internalized via clathrin-coated pits into vesicles. These vesicles very quickly shed their clathrin coat to become early endosomes. Receptors can internalize either with or without βArrestin. Receptors that internalize with βArrestin (e.g., β2-AR) are dephosphorylated in an acidified perinuclear compartment and then recycled back to the plasma membrane where they can once again interact with high-affinity ligands (resensitization) [45,46]. Receptors that internalize still bound to βArrestin (e.g., $AT_{1A}R$) are either retained in large endosomes and/or are targeted for degradation by lysosomes [47].

itization and internalization of certain GPCRs such as the β2-AR.

4. Signalling/scaffolding

In addition to the roles of βArrestin in linking phosphorylated receptors to the endocytic machinery, their roles in signalling are just beginning to be appreciated. βArrestin has been described as a molecular adapter linking phosphorylated receptors to clathrin and therefore targeting the receptors to clathrin-coated pits. Adapter proteins are best described as molecules that direct specific intermolecular interactions to orchestrate assembly of signalling complexes. Recently, βArrestins have been shown to play a critical role in coordinating the formation of signalling complexes following GPCR stimulation that ultimately leads to activation of certain members of the mitogen-activated protein kinase family (MAPKs) [26,27]. The core unit of a MAPK signalling complex comprises a three-tiered protein kinase cascade. Activation of each MAPK requires the dual phosphorylation of tyrosine and threonine residues by an upstream protein kinase (MAPKK) which in turn is activated by another protein kinase (MAPKKK) [28]. βArrestins have recently been demonstrated to regulate two MAPK family members, the extracellular signal regulated kinases (ERK1 and 2) [26] and the c-Jun NH_2-terminal kinase (JNK3) [27]. Activation of ERKs generally leads to their translocation to the nucleus where they activate numerous transcription factors that regulate cell growth and differentiation. Similarly, active JNK usually translocates to the nucleus where it interacts with a distinct subset of transcription factors generally associated with the stress response [28].

Many heptahelical and tyrosine kinase growth factor receptors have been shown to activate the extracellular signal-regulated kinase (ERK1/2) cascade, and, in some instances, clathrin-mediated endocytosis is required. In fact, initial evidence suggesting that βArrestins play a key role in ERK activation came from a study demonstrating that dominant negative forms of βArrestin1 that blocked endocytosis of the β2-AR-attenuated isoproterenol stimulated ERK activation [29]. The primary event that allows signal transmission to proceed from the plasma membrane bound receptor to ERK in the cytosol, is the assembly of a multiprotein complex at the plasma membrane that includes c-Src, Shc, Grb2, Gab1 and Sos. Assembly of this complex leads to the activation of the small G-protein Ras. Ras activation leads to stimulation of ERK, proceeding through the sequential phosphorylation and activation of the MAPKKK, Raf-1, and the MAPKK, MEK1.

As mentioned above, the Ras activation complex includes c-Src, a member of the Src family of tyrosine kinases. It has recently been demonstrated that stimulation of β2-ARs results in the assembly of a protein complex containing activated c-Src and the receptor [30]. c-Src recruitment can be mediated by βArrestin1, providing another example of βArrestin's ability to function as an adapter protein, binding both c-Src and agonist-occupied receptor. βArrestin1 not only serves to recruit c-Src to the agonist-occupied receptor but also targets the complex to clathrin-coated pits for internalization. Both events are required for ERK activation, as dominant negative forms of c-Src are effective in blocking this pathway [31,32]. In a similar manner, stimulation of the neurokinin-1 receptor (NK1R) by the neuropeptide substance P (SP), leads to activation of ERK1 and 2 through the formation of a signalling complex that contains internalized receptor, βArrestin1 and c-Src. Recruitment of c-Src to this complex is also βArrestin1 dependent [33]. Substance-P-stimulated ERK activation was found to be inhibited by expression of dominant negative βArrestin. Activation of ERK1/2 in this context leads to proliferation and protects the cells from apoptosis. Thus, βArrestin–c-Src complex formation was found to be required for the proliferative and antiapoptopic effects of the substance P.

More recent findings suggest that other members of the ERK activation pathway such as the MAPKKK Raf-1 and ERK itself interact with βArrestin. In the case of the proteinase-activated receptor 2 (PAR2), βArrestin forms a complex with the internalized receptor, Raf-1 and ERK1/2. Complex formation leads to ERK activation. In this case, activated ERK is retained in the cytosol and does not translocate to the nucleus. This inhibits the proliferative effects generally associated with translocation of the active kinase to the nucleus. However, retention of ERK in the cytosol may lead to phosphorylation of cytosolic rather than nuclear targets.

In addition, βArrestin has been shown to be involved in the activation and targeting of ERK2 following stimulation of the $AT_{1A}R$ [34]. Angiotensin treatment of cells leads to redistribution of activated ERK2 into endocytic vesicles that also contain $AT_{1A}R$–βArrestin complexes. This targeting of ERK2 reflects the formation of multiprotein complexes containing $AT_{1A}R$, βArrestin2 and the component kinases of the ERK cascade, cRaf-1, MEK and ERK2. The phosphorylation of ERK2 in βArrestin complexes was markedly enhanced by coexpression of cRaf-1, and this effect was found to be blocked by expression of a catalytically inactive mutant of MEK1. This data further supports the ability of the βArrestins to function both as scaffolds to enhance MAPKKK-dependent activation of MAPKs, and as targeting proteins that direct the activated MAPKs to specific subcellular locations.

It is clear that the biological consequences of ERK (MAPK) activation depend on the formation of specific signalling complexes, which are dictated by the original stimulus. Therefore, βArrestin complexes that differ in the receptor they contain but still lead to ERK activation might mediate distinct sets of cellular responses. In this regard, βArrestin acts as a molecular scaffold in which the functional integrity of each MAP kinase cascade is established

by specific molecular interactions both between the kinases and with βArrestin. These interactions result in the formation of multienzyme signalling complexes in which βArrestin localizes its binding partners to specific subcellular compartments or to specific substrates.

Scaffolding proteins that organize specific MAPK modules were first described in budding yeast. STE5 is a scaffold protein that forms a multicomponent complex with the Fus3 (Kss1) MAPK, Ste7 MAPKK and Ste11 MAPKKK to facilitate the specific and efficient activation of the mating pheromone pathway. Also in yeast, Pbs2, which itself is a MAPKK, has been proposed to serve as a scaffold protein in the HOG (high-osmolarity glycerol response) signal transduction pathway [28]. In mammalian cells, the MP1 protein has been shown to mediate activation of the ERKs [35] and a group of JNK interacting proteins (JIP1, 2, 3 and JSAP1) have been identified as scaffolds for specific JNK signalling modules [36–39]. For example, JIP1 was originally identified as a cytoplasmic protein that directly interacts with JNK1 and in doing so led to cytoplasmic retention of JNK, thus, preventing gene expression mediated by the JNK signalling pathway. In addition, JIP1 appeared to suppress the effect of JNK on apoptosis and malignant transformation [36]. However, these observations were based on an overexpressed phenotype. It is now believed that the function of JIP is to enhance JNK signalling. Thus, depending on the level of expression, the scaffold protein can either inhibit (excess scaffold) or enhance signalling (optimal expression) [40]. Following this, Whitmarsh et al. [37] demonstrated that JIP1 could interact with additional components of the JNK signalling module, including MKK7 (MAPKK) and MLK3 (MAPKKK). Thus, the function of these JIP scaffolds appears to be selective for the signalling pathway that is formed by the MAP kinase module MLK → MKK7 → JNK.

Recently, in a yeast two-hybrid screen, JNK3 was identified as a βArrestin2 binding partner. βArrestin2 was also found to interact with JNK3 and MAPKKK, ASK1 (apoptosis signalling kinase1) in cells. In addition, βArrestin2 was shown to act as a 'scaffold' molecule for the JNK3 MAPK cascade whereby overexpression of βArrestin2 in cells markedly enhances ASK1 stimulation of JNK3 via the MAPKK MKK4. Although MKK4 does not bind directly to βArrestin2, it appears to be recruited to the complex via ASK1 and/or JNK3 [27]. Fig. 2 is a graphical representation of the ASK1 mediated βArrestin2-dependent activation of JNK3. Furthermore, coexpression of βArrestin2 and JNK3 was found to cause cytosolic retention of JNK3. However, following agonist stimulation of the Angiotensin II receptor ($AT_{1A}R$), not only does stimulation of JNK3 phosphorylation occur but βArrestin2 and active JNK3 are found to colocalize in vesicles inside the cell.

Effectively, stimulation of the heptahelical receptor appears to regulate the subcellular distribution and activation of JNK3 by controlling its association with a cytoplasmic retention molecule, βArrestin2. Therefore, scaffold

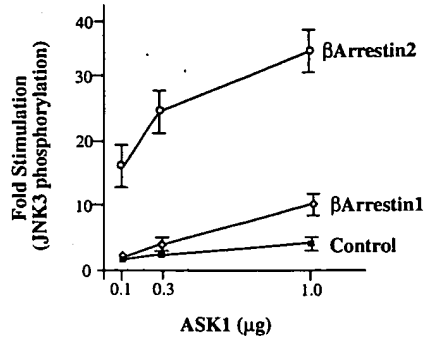


Fig. 2. βArrestin2 enhances ASK activation of JNK3. Whole cell lysates were prepared from cells expressing HA-JNK3 with either βArrestin1 or βArrestin2 or empty vector (Control) in the presence of increasing amounts of ASK1. The relative amounts of ASK1 and the βArrestins in the lysates was examined by protein blot analysis [25]. The results of four experiments are represented graphically, demonstrating that overexpression of βArrestin2 leads to a significant enhancement in ASK1 stimulated JNK3 activation.

proteins such as βArrestin2 appear to play a dual regulatory role. Firstly, they ensure precise regulation of the MAPK pathway by colocalizng and facilitating phosphorylation of successive members of the cascade. Secondly, they protect or insulate the MAPK pathway by preventing crosstalk with functionally unrelated members of other kinase signalling modules.

The idea that internalized receptor–arrestin complexes can trigger additional signalling events such as MAP kinase activation has recently been extended to include apoptosis [41,42]. βArrestin1 acts to desensitize the β2-AR by disrupting the G-protein–receptor interaction. Visual arrestin acts analogously to inactivate the phototransduction cascade by binding to light-activated phosphorylated rhodopsin and thereby uncoupling the receptor from its G-protein transducin. It has previously been demonstrated that the major *Drosophila* arrestin protein (Arr2) is phosphorylated very rapidly in a light-dependent manner [43] and that this phosphorylation event is essential for its release from rhodopsin containing complexes [44]. Therefore, a loss-of-function mutation that results in decreased phosphorylation of arrestin, results in the formation of stable rhodopsin–arrestin complexes. Furthermore, as visual arrestin is phosphorylated by the calcium/calmodulin-dependent protein kinase II, any mutations that result in lower levels of light-induced intracellular calcium should also yield stable interactions between rhodopsin and arrestin.

Using a combination of genetic and biochemical techniques, Alloway et al. [41] and Kiselev et al. [42] demonstrated that invertebrate Arrestin2 acts as a clathrin adapter protein and mediates internalization of rhodopsin–arrestin complexes. They further demonstrate that the long-term accumulation of rhodopsin–arrestin complexes in endocytic vesicles is the initiating factor necessary for photoreceptor cell apoptosis. In two distinct mutations, *norpA* (phospholipase C mutation) and *rdgB* (Ca^{2+}-dependent serine/threo-

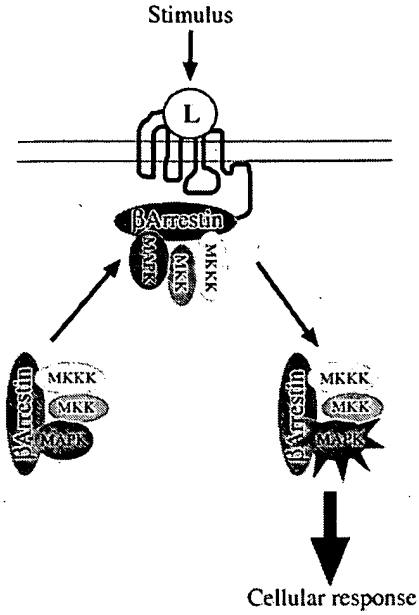


Fig. 3. βArrestin acts as a scaffold for MAPK modules. A complex is formed between βArrestin and the various components of a MAPK module in the cytosol. This multiprotein complex translocates to the plasma membrane bound receptor following ligand binding. βArrestin, acting as an adaptor, promotes internalization of the whole complex via clathrin-coated pits. This leads to activation of the MAPK module and, hence, generation of the active MAPK.

nine protein phosphatase) light-dependent formation of stable rhodopsin–arrestin complexes triggers apoptopic cell death. Elimination of either member of the rhodopsin–arrestin complex rescues the retinal degeneration phenotype. It has been proposed that endocytosis of these complexes and their prolonged presence in the cytosol is required for the initiation of apoptosis of the retinal photoreceptors and that excessive endocytosis, and, therefore, constitutive signalling, is the direct cause of apoptopic retinal degeneration. The arrestin-dependent signalling pathway(s) leading to the apoptosis however have not been clarified.

5. Conclusions

βArrestins are emerging as multifaceted proteins with functions quite distinct from their well-defined roles in desensitizing and sequestering agonist-occupied receptors. In addition to desensitizing heptahelical receptors, βArrestins have been shown to function as GPCR signalling molecules, scaffolds and adapters. Therefore, βArrestin binding not only terminates receptor–G-protein coupling but also initiates a second wave of signal transduction in which it functions as a critical structural scaffold linking the 'desensitized' receptor to MAP kinase signalling modules (Fig. 3). The growing list of functions described here suggests that our appreciation of the number of functions served by arrestins will most likely continue to expand.

Acknowledgments

We would like to thank Drs. KP Pierce and WE Miller for critical reading of the manuscript. RJL is an investigator of the Howard Hughes Medical Institute.

References

[1] Pitcher JA, Freedman NJ, Lefkowitz RJ. Annu Rev Biochem 1998;67:653–92.
[2] Ferguson SSG, Barak LS, Zhang J, Caron MG. Can J Physiol Pharmacol 1996;74:1095–110.
[3] Sterne-Marr R, Benovic JL. Vitam Horm 1995;51:193–235.
[4] Krupnick JG, Benovic JL. Annu Rev Pharmacol Toxicol 1998;38:289–319.
[5] Pfister C, Habre M, Plouet J, Tuyen VV, De Kozak Y, Faure JP, Kuhn H. Science (Washington, D. C.) 1985;228:891–3.
[6] Lohse MJ, Benovic JL, Codina J, Caron MJ, Lefkowitz RJ. Science 1990;248:1547–50.
[7] Attramadal H, Arriza JL, Aoki C, Dawson TM, Codina J, Kwatra MM, Snyder SH, Caron MG, Lefkowitz RJ. Proc Natl Acad Sci USA 1992;267:17882–90.
[8] Ferguson SSG, Downey WE, Colapietro A-M, Barak LS, Menard L, Caron MG. Science 1996;271:363–6.
[9] Goodman OB, Krupnick JG, Santini F, Gurevich VV, Penn RB, Gagnon AW, Keen JH, Benovic JL. Nature 1996;383:447–50.
[10] Laporte SA, Oakley RH, Zhang J, Holt JA, Ferguson SSG, Caron MG, Barak LS. Proc Natl Acad Sci USA 1999;96:3712–7.
[11] McDonald PH, Cote NL, Lin F-T, Premont RT, Pitcher JA, Lefkowitz RJ. J Biol Chem 1999;274:10677–81.
[12] Pippig S, Andexinger S, Daniel K, Puzicha M, Caron MG, Lefkowitz RJ, Lohse M. J Biol Chem 1993;268:3201–8.
[13] Freedman NJ, Ligget SB, Drachman DE, Pei G, Caron MG, Lefkowitz RJ. J Biol Chem 1995;270:17953–61.
[14] Dawson TM, Arriza JL, Jaworsky DE, Borisy FF, Attramadal H, Lefkowitz RJ, Ronnett GV. Science 1993;259:825–9.
[15] Dolph PJ, Ranganathan R, Colley NJ, Hardy RW, Socolich M, Zucker CS. Science 1993;260:1910–6.
[16] Conner DA, Mathier MA, Mortensen RM, Christe M, Vatner SF, Seidman CE, Seidman JG. Circ Res 1997;81:1021–6.
[17] Bohn LM, Lefkowitz RJ, Gainetdinov RR, Peppel K, Caron MG, Lin F-T. Science 1999;286:2495–8.
[18] Bohn LM, Gainetdinov RR, Lin F-T, Lefkowitz RJ, Caron MG. Nature 2000;408:720–3.
[19] Lefkowitz RJ. Cell 1993;74:409–12.
[20] Goodman OB, Krupnick JG, Gurevich VV, Benovic JL, Kenn JH. J Biol Chem 1997;272:15017–22.
[21] Krupnick JG, Santini F, Gagnon AW, Keen JH, Benovic JL. J Biol Chem 1997;272:32507–12.
[22] Krupnick JG, Goodman OB, Keen JH, Benovic JL. J Biol Chem 1997;272:15011–6.
[23] Ferguson SSG, Menard L, Barak LS, Colapietro A-M, Koch WJ, Caron MG. J Biol Chem 1995;270:24782–9.
[24] Lin F-T, Kreuger KM, Kendall HE, Daaka Y, Fredericks ZL, Pitcher JA, Lefkowitz RJ. J Biol Chem 1997;272:31051–7.
[25] Laporte SA, Oakley RH, Holt JA, Barak LS, Caron MG. J Biol Chem 2000;275:23120–6.
[26] Defea KT, Zalevesky J, Thoma MS, Dery O, Mullins RD, Bunnet NW. J Cell Biol 2000;148:1267–81.
[27] McDonald PH, Chow C-W, Miller WE, Laporte SA, Field ME, Lin F-T, Davis RJ, Lefkowitz RJ. Science 2000;290:1574–7.
[28] Whitmarsh AJ, Davis RJ. Trends Biochem Sci 1998;23:481–5.
[29] Daaka Y, Luttrell LM, Ahn S, Della Rocca GJ, Ferguson SSG, Caron MG, Lefkowitz RJ. J Biol Chem 1998;273:685–8.

[30] Luttrell LM, Ferguson SSG, Daaka Y, Miller WE, Maudsley S, Della Rocca GT, Lin F-T, Kawakatsu H, Owada K, Luttrell DK, Caron MG, Lefkowitz RJ. Science 1999;283:655–61.
[31] Maudsley S, Pierce KL, Zamah AM, Miller WE, Ahn S, Daaka Y, Lefkowitz RJ, Luttrell LM. J Biol Chem 2000;275:9572–80.
[32] Miller WE, Mausdley S, Ahn S, Khan K-D, Luttrell LM, Lefkowitz RJ. J Biol Chem 2000;275:11321–9.
[33] Defea KA, Vaughn ZD, O'Bryan ME, Nishijima D, Dery O, Bunnett NW. Proc Natl Acad Sci USA 2000;97:11086–91.
[34] Luttrell LM, Roudabush FL, Choy EW, Miller WE, Field ME, Pierce F-T, Lin F-T, Lefkowitz RJ. Proc Natl Acad Sci USA 2001;98: 2449–54.
[35] Schaeffer HJ, Catling AD, Eblen ST, Collier LS, Krauss A, Weber MJ. Science 1998;281:1668–71.
[36] Dickens M, Rogers RS, Cavanagh J, Sharma M, Davis RJ. Science 1997;277:693–6.
[37] Whitmarsh AJ, Cavanagh J, Tournier C, Yusada J, Davis RJ. Science 1998;281:1671–4.
[38] Yasuda J, Whitmarsh AJ, Cavanagh J, Sharma M, Davis RJ. Mol Cell Biol 1999;19:7245–54.
[39] Ito M, Yoshioka K, Akechi M, Yamashita S, Takamatsu N, Sugiyama K, Hibi M, Nakabeppu Y, Shiba T, Yamamoto KI. Mol Cell Biol 1999;19: 7539–48.
[40] Levchenko A, Bruck J, Sternberg PW. Proc Natl Acad Sci USA 2000;97:5818–23.
[41] Alloway PG, Howard L, Dolph PJ. Neuron 2000;28:129–38.
[42] Kiselev A, Socolich M, Vinos J, Hardy RW, Zuker CS, Ranganathan R. Neuron 2000;28:139–52.
[43] Kahn ES, Matsumoto H. J Neurochem 1997;68:169–75.
[44] Alloway PG, Dolph PJ. Proc Natl Acad Sci USA 1999;96:6072–7.
[45] Oakley RH, Laporte SA, Holt JA, Barak LS, Caron MG. J Biol Chem 1999;274:32248–57.
[46] Kreuger KM, Daaka Y, Pitcher JA, Lefkowitz RJ. J Biol Chem 1997;272:5–8.
[47] Anborgh PH, Dale L, Seachrist J, Ferguson SSG. Mol Endocrinol 2000;14:2040–53.

Appendix II

Communication

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 272, No. 44, Issue of October 31, pp. 27497–27500, 1997
© 1997 by The American Society for Biochemistry and Molecular Biology, Inc.
Printed in U.S.A.

A β-Arrestin/Green Fluorescent Protein Biosensor for Detecting G Protein-coupled Receptor Activation*

(Received for publication, August 21, 1997)

Larry S. Barak, Stephen S. G. Ferguson‡, Jie Zhang, and Marc G. Caron§

From the Howard Hughes Medical Institute Laboratories and Department of Cell Biology, Duke University Medical Center, Durham, North Carolina 27710

G protein-coupled receptors (GPCR) represent the single most important drug targets for medical therapy, and information from genome sequencing and genomic data bases has substantially accelerated their discovery. The lack of a systematic approach either to identify the function of a new GPCR or to associate it with a cognate ligand has added to the growing number of orphan receptors. In this work we provide a novel approach to this problem using a β-arrestin2/green fluorescent protein conjugate (βarr2-GFP). It provides a real-time and single cell based assay to monitor GPCR activation and GPCR-G protein-coupled receptor kinase or GPCR-arrestin interactions. Confocal microscopy demonstrates the translocation of βarr2-GFP to more than 15 different ligand-activated GPCRs. These data clearly support the common hypothesis that the β-arrestin binding of an activated receptor is a convergent step of GPCR signaling, increase by 5-fold the number of GPCRs known to interact with β-arrestins, demonstrate that the cytosol is the predominant reservoir of biologically active β-arrestins, and provide the first direct demonstration of the critical importance of G protein-coupled receptor kinase phosphorylation to the biological regulation of β-arrestin activity and GPCR signal transduction in living cells. The use of βarr2-GFP as a biosensor to recognize the activation of pharmacologically distinct GPCRs should accelerate the identification of orphan receptors and permit the optical study of their signal transduction biology intractable to ordinary biochemical methods.

The G protein-coupled receptor (GPCR)[1] superfamily is growing rapidly (1–3), creating many new orphan receptors whose properties remain undefined (4–6).[2] Currently characterized GPCRs display many distinct pharmacologies. For example, they interact with a vast array of ligands and generate intracellular signals by multiple second messenger pathways (4, 8, 9). Based on work with rhodopsin and the β_2-adrenergic receptor (β_2AR), it has been postulated that members of the GPCR superfamily desensitize via a common mechanism involving the arresting proteins visual arrestin, β-arrestin1 and β-arrestin2 (10–13). However, mainly due to the inherent difficulties of examining the interaction of the components mediating desensitization in their native environment or the need for purified reconstituted systems, this has not been clearly established for many GPCRs. Biochemical studies indicate that arrestins regulate GPCR signal transduction (desensitization) by binding agonist-activated receptors that have been phosphorylated by G protein-coupled receptor kinases (GRKs) (12). While the functional source of arrestin molecules targeted to receptors remains unknown, it is apparent that arrestin binding terminates signaling by interdicting receptor interaction with G proteins (12).

To characterize the interaction between β-arrestin and different GPCRs and to assess the contribution of GRKs to this process, we examined using confocal microscopy how a green fluorescent protein/β-arrestin2 conjugate responded to ligand-mediated receptor activation. The results demonstrate a critical role for GRKs in the apparently universal regulation of GPCRs by β-arrestins. Moreover, they provide the first real-time, live-cell demonstration of a GPCR interacting with one of its regulatory proteins and demonstrate a practical role for βarr2-GFP in the study of GPCR activation.

EXPERIMENTAL PROCEDURES

Materials—Isoproterenol was obtained from Sigma and Research Biochemicals International. Anti-mouse antibody was obtained from Sigma and Molecular Probes. Mouse monoclonal antibody against the 12CA5 epitope was purchased from Boehringer Mannheim. Cell culture medium was obtained from Mediatech and fetal bovine serum from Atlanta Biologicals. Physiological buffers were from Life Technologies, Inc. Restriction enzymes were obtained from Promega or New England Biolabs, T4 ligase was from Promega, and Hot Tub DNA polymerase from Amersham. Plasmid containing variants of green fluorescent protein and anti-GFP antibodies were from CLONTECH.

Plasmid Construction, βarr2-GFP—Oligonucleotide primers surrounding the *Xho*I restriction site and C-terminal stop codon of β-arrestin2, in the expression vector pCMV5, were used to replace the stop codon with an in frame *Bam*HI restriction site by directed mutagenesis (14). The product nucleotide was purified by electrophoresis on a 1.5% agarose gel and isolated. It was digested using *Xho*I and *Bam*HI restriction enzymes, and the nucleotide fragment flanked by these sites was repurified. The N-terminal, proximal cDNA fragment of β-arrestin2 flanked by *Sac*I/*Xho*I restriction sites was removed by digestion from pCMV5. It was ligated with the *Xho*I/*Bam*HI fragment isolated above and with the purified expression vector (pS65T-GFP) that had been opened between the *Sac*I/*Bam*HI polylinker restriction sites using the respective enzymes (15). The resulting construct was grown in competent *Escherichia coli*, isolated, and verified by sequencing.

Cell Culture and Transfection—HEK-293 and COS cells were maintained and transfected as described previously (15, 16). Cells containing both receptor and β-arrestin constructs were transfected with between 5 and 10 μg of receptor cDNA in pcDNA1/AMP and 0.5–1 μg of βarr2-

* This work was supported by National Institutes of Health Grant NS 19576, a Bristol Myers Squibb Unrestricted Grant Award, and an unrestricted grant from Zeneca Pharmaceuticals (to M. G. C.) and National Institutes of Health Grant HL 03422 (to L. S. B.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Recipient of a Michael Smith Postdoctoral Fellowship from the Medical Research Council, Canada. Present address: The John P. Robarts Inst., P. O. Box 5015, 100 Perth Dr., London, Ontario N6A 5K8, Canada.

§ To whom correspondence should be addressed: Duke University Medical Center, Box 3287, Durham, NC 27710. Tel.: 919-684-5433; Fax: 919-681-8641; E-mail: caron002@mc.duke.edu.

[1] The abbreviations used are: GPCR, G protein-coupled receptor; GFP, green fluorescent protein; βarr2-GFP, β-arrestin2 green fluorescent protein conjugate; GRK, G protein-coupled receptor kinase; β_2AR, β_2-adrenergic receptor.

[2] On the World Wide Web at receptor.mgh.harvard.edu/GCRDHOME.html.

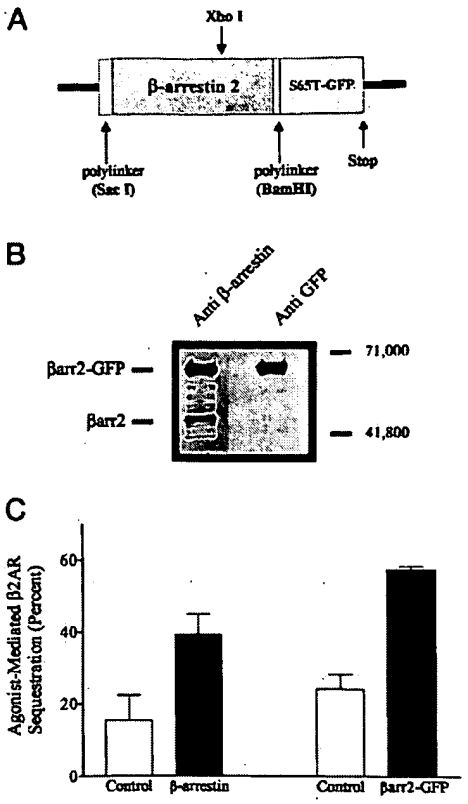


FIG. 1. **Linear model of the β-arrestin2/S65T-GFP conjugate and its characterization by SDS-polyacrylamide gel electrophoresis and β2AR sequestration.** *A*, the model indicates the relative size of β-arrestin2 compared with GFP. In this construct the C-terminal stop codon of β-arrestin2 is replaced by a *Bam*HI restriction site, and β-arrestin2 cDNA is inserted in frame to GFP between *Sac*I and *Bam*HI. *B*, shown are Western blots of homogenates from HEK-293 cells expressing βarr2-GFP. Equal amounts of material were loaded into each lane. The gel on the *left* was exposed to anti-β-arrestin antibody (16), whereas the *right gel* was exposed to a mouse monoclonal antibody against GFP. The position of endogenous cellular β-arrestin2 is indicated by the *lower bar* on the *left*. The heavy band below 71,000 (*upper left*) corresponding to βarr2-GFP is mirrored by a similar band in the *right gel*. In contrast, no band corresponding to endogenous cellular β-arrestin2 is observed with anti-GFP antibody exposure. *C*, β-arrestin activity can indirectly be assessed by measuring its effect on receptor sequestration (16, 22). The sequestration of β_2AR in COS cells with and without overexpressed β-arrestin2 (*left bars*) or with and without overexpressed βarr2-GFP (*rightmost bars*) is presented. Wild type β-arrestin2 and βarr2-GFP enhance β2AR sequestration equally well, producing a 2.5- and 2.4-fold increase, respectively. (Results are average ± S.D. from triplicate experiments.)

GFP cDNA/100-mm dish. GRKs were expressed using 5 μg of transfected cDNA in pcDNA1/AMP per dish.

Confocal Microscopy—HEK-293 cells transfected as described above were plated onto 35-mm dishes containing a centered, 1-cm well formed from a glass coverslip sealed hole in the plastic. Primary and secondary antibody labeling of live cells was performed at 37 °C for 30 min in media without serum in a 5% CO_2 incubator. Cells were washed three times between applications. Cells plated as above in minimal essential medium or Dulbecco's modified Eagle's medium buffered with 20 mM Hepes were viewed on a Zeiss laser scanning confocal microscope.

Sequestration—Flow cytometry analysis was performed as described (17).

RESULTS

GFP, because of its inherent fluorescence, represents a valuable biological reporter molecule for the study of GPCR signal transduction events (15, 18–21). The βarr2-GFP fusion protein (Fig. 1*A*), which is approximately 50% larger than β-arrestin2 and migrates more slowly on SDS-polyacrylamide gel electrophoresis (Fig. 1*B*), still retains its biological activity with respect to facilitating β_2AR sequestration (22). In the absence of supplemental β-arrestins, the β_2AR normally sequesters poorly in COS-7 cells (16). βarr2-GFP overexpression increases sequestration to the same extent as wild type β-arrestin2. (Fig. 1*C*).

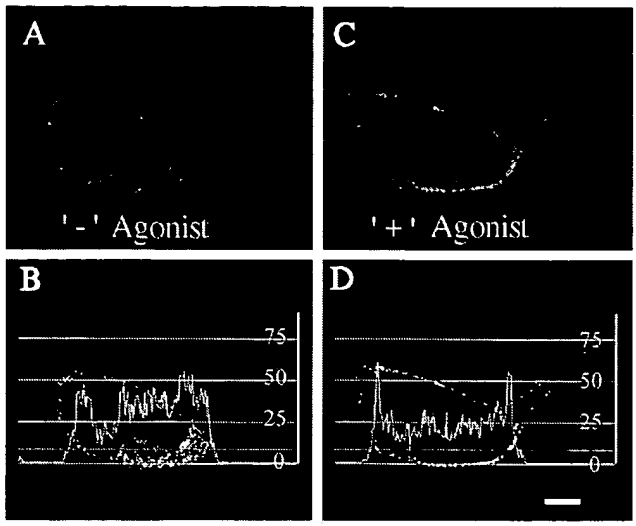


FIG. 2. **Confocal Microscopy of βarr2-GFP translocation in cells containing the β2AR.** HEK-293 cells were transfected with 10 μg of plasmid cDNA containing the β_2AR and 0.5–1.0 μg for βarr2-GFP. *A* and *B*, βarr2-GFP distribution is initially cytosolic with no significant nuclear or membrane enhancement apparent. *C* and *D*, following a 10 min stimulation with 20 μM isoproterenol, the real-time agonist-mediated redistribution of βarr2-GFP from the cytosol to the membrane is shown. The relative magnitudes of the βarr2-GFP concentrations can be quantified (*A*) before and (*C*) after agonist treatment. βarr2-GFP concentrations along a linear slice (*blue line*) made through the lower portion of the above cell in *A* and *C* are profiled in *B* and *D*, respectively. The flatness of the central portion of the trace in *B* corresponds to the uniform cytoplasmic distribution of βarr2-GFP prior to agonist treatment. However, in regions lacking βarr2-GFP, *i.e.* outside the cell and in the nucleus, the magnitude of the trace decreases. In contrast, the concave trace obtained after isoproterenol treatment, (*D*), reflects the increased concentration of βarr2-GFP at the cell margins and the reduction of βarr2-GFP concentration in the cell cytoplasm that results from GPCR activation by agonist. *Bar* = 10 μm.

Confocal microscopy of βarr2-GFP in an HEK-293 cell (Fig. 2*A*) shows that in the absence of receptor activation β-arrestins are distributed throughout the cytosol and excluded from the nucleus. Moreover, the data demonstrate that β-arrestins are not predominantly compartmentalized at the plasma membrane prior to agonist stimulation. Upon the addition of saturating concentrations of the agonist isoproterenol to the cell medium, an enhancement of plasma membrane fluorescence and a concomitant loss of cytosolic fluorescence (Fig. 2*B*) can be readily observed and quantified (Figs. 2, *C* and *D*). This observation indicates that β-arrestins are not discretely compartmentalized and that the entire cytoplasmic content represents a functional β-arrestin reservoir.

HEK-293 cells overexpressing the β_2AR were used to investigate whether the main target of translocated βarr2-GFP was a plasma membrane site other than a GPCR. N-terminal epitope-tagged β_2ARs were cross-linked to one another prior to agonist exposure using a mouse monoclonal antibody against the epitope and a secondary goat anti-mouse antibody conjugated to the fluorophore Texas Red. Fig. 3 demonstrates that the geometry of the agonist-induced time-dependent translocation of βarr2-GFP to the plasma membrane mirrors the distribution of preaggregated β2ARs, strongly suggesting that the primary targeted site of β-arrestin is the β_2AR.

βarr2-GFP translocation to β2ARs is not limited to HEK-293

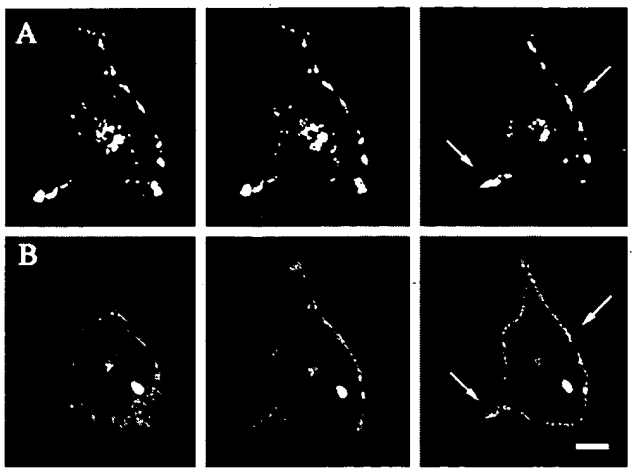


FIG. 3. **Time course of βarr2-GFP redistribution to plasma membrane, 12CA5 (hemagglutinin)-tagged β_2AR in HEK-293 cells as observed by confocal microscopy.** To demonstrate that β_2ARs were the target of intracellular βarr2-GFP, receptors were reorganized into plasma membrane clusters (*A*) by cross-linking with a primary layer of monoclonal anti-epitope antibody followed by a secondary layer of Texas Red-conjugated goat anti-mouse antibody. The agonist-mediated time course of βarr2-GFP redistribution to these receptors at 0, 3, and 10 min (*B*, *panels left* to *right*) is shown to correspond to the aggregated receptor distribution in (*A*, corresponding *panels left* to *right*). *Arrows* indicate some of the regions of colocalization. *Bar* = 10 μm.

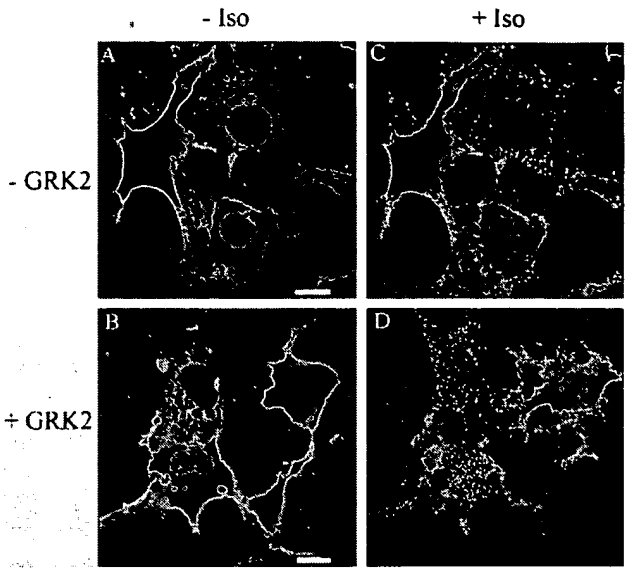


FIG. 4. **Influence of overexpressed GRK on redistribution of βarr2-GFP in COS-7 cells.** *A–D*, COS-7 cells were transiently transfected with βarr2-GFP and β2AR as described, and the agonist-mediated redistribution of βarr2-GFP in the absence (*A*, *C*) or presence (*B*, *D*) of overexpressed GRK2 was determined. Both the rate and extent of agonist-mediated βarr2-GFP translocation are observed to be enhanced by the overexpression of GRKs, and this is reflected by an increase in contrast in the photomicrographs between cytoplasm and membrane βarr2-GFP fluorescence for the GRK untreated pair (*A*, *C*) and the GRK-treated pair (*B*, *D*). *Bar* = 10 μm.

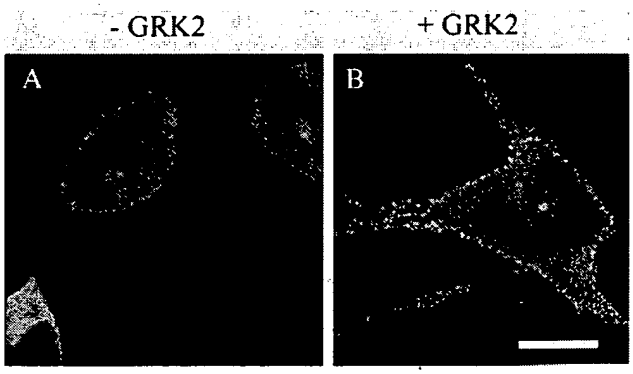


FIG. 5. **Influence of overexpressed GRK on redistribution of βarr2-GFP to the Y326A-β_2AR in HEK-293 cells.** The effect of overexpressed GRK2 on the agonist-mediated interaction between the phosphorylation-impaired Y326A-β_2AR and βarr2-GFP in HEK-293 cells is shown. Cells without (*A*) and with (*B*) overexpressed GRK2 were exposed to agonist. βarr2-GFP translocation in cells containing overexpressed GRK2 is more robust, as demonstrated (*B*), indicating an increased affinity of βarr2-GFP for receptor. *Bar* = 10 μm.

cells, but is also observable in COS-7 cells (Fig. 4). Consistent with their relatively larger surface area and lower efficiency of β_2AR sequestration compared with HEK-293 cells, agonist-mediated βarr2-GFP translocation was less apparent in the COS cells. (Fig. 4, *A* and corresponding image *C*) (16). However, by coexpressing GRK2, the agonist-mediated βarr2-GFP translocation could be enhanced (Fig. 4, *B* and corresponding image *D*), suggesting that GRK phosphorylation increases the affinity of the receptor for β-arrestin (22). To further characterize the role of GRK phosphorylation in β-arrestin translocation, we examined the ability of a GRK phosphorylation-impaired mutant Y326A-β_2AR to support β-arrestin redistribution (11, 17, 23, 24). Consistent with its inability to be phosphorylated by endogenous GRKs, the Y326A-β2AR mutant did not induce βarr2-GFP translocation with agonist exposure (Fig. 5*A*). However, with overexpression of GRK2 and agonist treatment, the Y326A mutant-mediated βarr2-GFP translocation (Fig. 5*B*) was indistinguishable from β_2AR-mediated translocation (7, 25–26). These results indicate βarr2-GFP translocation not only accurately monitors the biology of the GPCR activation process but the GPCR phosphorylation state as well.

To establish that agonist-induced βarr2-GFP translocation represents a general property of GPCR activation and is not limited to the β2AR, other members of the GPCR superfamily were evaluated for their ability to mediate the movement of βarr2-GFP in HEK-293 cells. Shown in Fig. 6 are results with the dopamine D_{1A} receptor. Its behavior is representative of 16 different GPCRs that were tested belonging to the the angiotensin, α- and β-adrenergic, dopamine, endothelin, intestinal peptide, chemokine, and opioid receptor subfamilies. Activation of the D_{1A} receptor with 20 μM dopamine produced an increase in the amount of membrane-associated βarr2-GFP. Moreover, the increase in βarr2-GFP translocation was enhanced in the presence of overexpressed GRK2 (Fig. 6*D*).

DISCUSSION

In this work we demonstrate that β-arrestin interacts with GPCRs immediately following agonist stimulation and GRK phosphorylation. βarr2-GFP translocation was observed in response to more than 15 different GPCRs. Even though these GPCRs respond to a diverse array of ligands and different classes of G proteins, activation of each of the GPCRs elicits the agonist-dependent translocation of βarr2-GFP, with the magnitudes of plasma membrane fluorescence signals ranging up to 10–20-fold above the intracellular background. While β-arrestin behavior is regulated by multiple components of the signal transduction system, it is particularly sensitive to how well cellular GRKs are able to phosphorylate a particular GPCR. This was demonstrated with both GRK2 and GRK5 (data not shown). For instance, following overexpression of GRK2 to force phosphorylation of the Y326A-β_2AR mutant, the mutant-mediated βarr2-GFP translocation is indistinguishable from the wild type β_2AR-mediated response. Consequently,

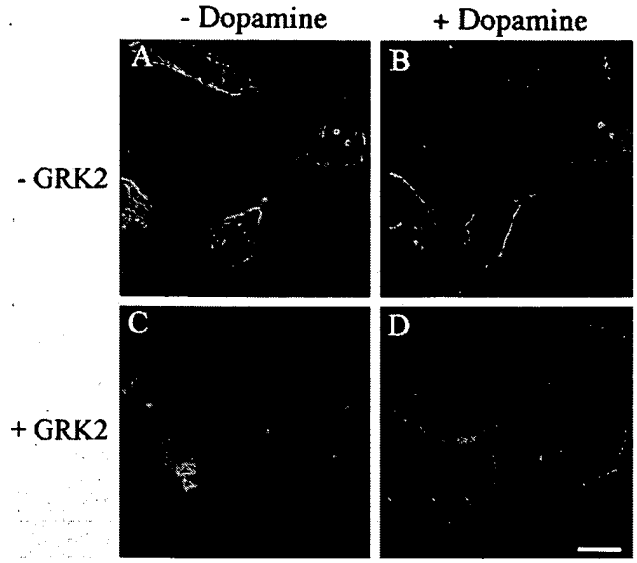


FIG. 6. **Agonist-mediated βarr2-GFP translocation to the dopamine D_{1A} receptor in HEK-293 cells.** *A–D*, the dopamine D_{1A} receptor and βarr2-GFP were expressed in HEK-293 cells in the absence and presence of GRK2 as described under "Experimental Procedures." Paired images of cells (*A* and *B*, *C* and *D*) are shown before and after agonist treatment. Prior to dopamine treatment (*A*, *C*) βarr2-GFP is found uniformly throughout the cytosol. Upon addition of 20 μM dopamine βarr2-GFP can be observed to translocate to the plasma membrane (*B*, *D*). The relative amount of βarr2-GFP translocated to the plasma membrane is enhanced by the presence of GRK2 (agonist-treated, GRK-untreated cell group B compared with agonist-treated, GRK2-treated cell group *D*). *Bar* = 10 μm.

with the appropriate cellular system, such as COS-7 cells in which endogenous GRKs and β-arrestins are relatively poorly expressed, the β-arrestin translocation paradigm could also be used to easily monitor the activity and specificity of each of the members of the GRK and arrestin families.

Biochemical measurements of GPCR properties, such as ligand binding, activation of G proteins or effectors, generation of second messengers, or extent of phosphorylation, assess functions that are receptor-specific and do not easily lend themselves to the development of rapid or convenient screening methods. However, since GPCR activation ultimately terminates with the association of β-arrestin and receptor, a convergent step of the GPCR signal transduction paradigm, the cellular visualization of the agonist-mediated translocation of βarr2-GFP provides a universal measure for detecting the activation of unknown GPCRs. Despite its present large size, the G protein-coupled receptor superfamily continues to expand rapidly as new receptors are discovered through automated sequencing of cDNA libraries. It is estimated that several thousand GPCRs may exist in the human genome, and at present with only a fraction of the genome sequenced, as many as 250 GPCRs have been cloned and only as few as 150 have been associated with ligands (4).[2] The means by which these or newly discovered orphan receptors will be associated with their cognate ligands and physiological functions represents a major challenge to biological and biomedical research. The identification of an orphan receptor generally requires an individualized assay and a guess as to its function. The interrogation of a GPCR's signaling behavior by monitoring βarr2-GFP translocation eliminates these prerequisites, since it can be performed with unlabeled ligands and without any prior knowledge of other signaling events. It is sensitive, rapid, easily performed, and should be potentially applicable to nearly all GPCRs, since the majority of these receptors should desensitize by a common mechanism, *i.e.* interaction with β-arrestins.

The visualization of β-arrestin2 translocation represents the first direct real-time assessment in a living cell of the interaction of a GPCR with one of its regulatory components. Moreover, the rapid and profound increases in the relative and absolute amounts of plasma membrane-bound β-arrestin provide an optical detection of GPCR signal transduction that is as sensitive as any chemical amplification normally produced by second messenger cascades. Therefore, βarr2-GFP is not only exquisitely adept as a biosensor for monitoring GPCR activation, but represents an excellent tool to study the kinetics and specificity of components involved in the regulation of GPCR activity. Furthermore, when used as an optical sensor, βarr2-GFP provides the unique potential to unite orphan GPCRs with their corresponding ligands.

REFERENCES

1. Hillier, L., Lennon, G., Becker, M., Bonaldo, M., Chiapelli, B., Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M., Tan, T., Thierry-Meg, J., Trevaskis, E., Underwood, K., Wohldman, P., Waterson, R., Wilson, R., and Marra, M. (1996) *Genome Res.* **6**, 807–828
2. Adams, M. D., Kelley, J. M., Gocayne, D., Dubnick, M., Polymeropoulos, M. H., Xiao, H., Merril, C. R., Wu, A., Olde, B., Moreno, R. F., Kerlavage, A. R., McCombie, W. R., and Venter, J. C. (1991) *Science* **252**, 1651–1656
3. Wells, T., and Peitsch, M. (1997) *J. Leukocyte Biol.* **61**, 545–550
4. Probst, W. C., Snyder, L. A., Schuster, D. I., Brosius, J., and Sealfon, S. C. (1992) *DNA Cell Biol.* **11**, 1–20
5. Oliver, S. (1996) *Nature* **379**, 597–600
6. Lopeznieto, C., and Nigam, S. (1996) *Nat. Biotechnol.* **14**, 857–861
7. Tsuga, H., Kameyama, K., Haga, T., Kurose, H., and Nagao, T. (1994) *J. Biol. Chem.* **269**, 522–527
8. Raymond, J. R., Hnatowich, M., Lefkowitz, R. J., and Caron, M. G. (1990) *Hypertension* **15**, 119–131
9. Ross, E. M. (1990) *Goodman and Gilman's The Pharmacological Basis of Therapeutics*, pp. 33–48, Pergamon Press, New York
10. Gurevich, V. V., and Benovic, J. L. (1995) *J. Biol. Chem.* **270**, 6010–6016
11. Gurevich, V. V., Richardson, R. M., Kim, C. M., Hosey, M. M., and Benovic, J. L. (1993) *J. Biol. Chem.* **268**, 16879–16882
12. Ferguson, S. S. G., Barak, L. S., Zhang, J., and Caron, M. G. (1996) *Can. J. Physiol. Pharmacol.* **74**, 1095–1110
13. Sterne-Marr, R., and Benovic, J. L. (1995) *Vitam. Horm.* **51**, 193–232
14. Valette, F., Mege, E., Reiss, A., and Adesnik, M. (1989) *Nucleic Acids Res.* **17**, 723–733
15. Barak, L. S., Ferguson, S. S. G., Zhang, J., Martenson, C., Meyer, T., and Caron, M. G. (1997) *Mol. Pharmacol.* **51**, 177–184
16. Menard, L., Ferguson, S. S. G., Zhang, J., Lin, F.-T., Lefkowitz, R. J., Caron, M. G., and Barak, L. S. (1997) *Mol. Pharmacol.* **51**, 800–808
17. Barak, L. S., Tiberi, M., Freedman, N. J., Kwatra, M. M., Lefkowitz, R. J., and Caron, M. G. (1994) *J. Biol. Chem.* **269**, 2790–2795
18. Kaether, C., and Gerdes, H. H. (1995) *FEBS Lett.* **369**, 267–271
19. Olson, K. R., Mcintosh, J. R., and Olmsted, J. B. (1995) *J. Cell Biol.* **130**, 639–650
20. Prasher, D. C., Eckenrode, V. K., Ward, W. W., Prendergast, F. G., and Cormier, M. J. (1992) *Gene (Amst.)* **111**, 229–233
21. Ormo, M., Cubitt, A. B., Kallio, K., Gross, L. A., Tsien, R. Y., and Remington, S. J. (1996) *Science* **273**, 1392–1395
22. Ferguson, S. S., Downey, W. E., III, Colapietro, A. M., Barak, L. S., Menard, L., and Caron, M. G. (1996) *Science* **271**, 363–366
23. Barak, L. S., Menard, L., Ferguson, S. S., Colapietro, A. M., and Caron, M. G. (1995) *Biochemistry* **34**, 15407–15414
24. Gurevich, V. V., and Benovic, J. L. (1993) *J. Biol. Chem.* **268**, 11628–11638
25. Menard, L., Ferguson, S. S., Barak, L. S., Bertrand, L., Premont, R. T., Colapietro, A. M., Lefkowitz, R. J., and Caron, M. G. (1996) *Biochemistry* **35**, 4155–4160
26. Ferguson, S. S., Menard, L., Barak, L. S., Koch, W. J., Colapietro, A. M., and Caron, M. G. (1995) *J. Biol. Chem.* **270**, 24782–24789

Appendix III

Review TRENDS in Pharmacological Sciences Vol.25 No.8 August 2004
Full text provided by www.sciencedirect.com
SCIENCE DIRECT®

Historical review: A brief history and personal retrospective of seven-transmembrane receptors

Robert J. Lefkowitz

Howard Hughes Medical Institute, James B. Duke Professor of Medicine, Duke University Medical Center, Room 468, CARL Bldg., Research Drive, DUMC Box 3821, Durham, NC 27710, USA

Pharmacologists have studied receptors for more than a century but a molecular understanding of their properties has emerged only during the past 30–35 years. In this article, I provide a personal retrospective of how developments and discoveries primarily during the 1970s and 1980s led to current concepts about the largest group of receptors, the superfamily of seven-transmembrane (7TM) receptors [also known as G-protein-coupled receptors (GPCRs)]. Significant technical advances such as the development of methods for radioligand binding, solubilization and purification of the β_2-adrenoceptor and other adrenoceptors led to the cloning of receptor genes and the discovery of their 7TM architecture and homology with rhodopsin. A universal mechanism of receptor regulation by G-protein-coupled receptor kinases (GRKs) and arrestins, originally discovered as a means of 'desensitizing' G-protein-mediated second-messenger generation, was subsequently found to mediate both receptor endocytosis and activation of a growing list of signaling pathways such as those involving mitogen-activated protein kinases. Numerous opportunities for novel therapeutics should emerge from current and future research on 7TM receptor biology.

Receptors for drugs, hormones and neurotransmitters have fascinated biologists for more than a century, and have fascinated me for the entirety of my 35-year research career. This fascination continues to engage a significant number of biomedical scientists. For example, PubMed lists >27 000 articles published in 2003 alone on the topic of 'receptors'.

Of the several large families of receptors, by far the largest, most versatile and most ubiquitous is that of the seven-transmembrane (7TM) receptors [also referred to as seven-membrane-spanning receptors and G-protein-coupled receptors (GPCRs)] [1]. There are ~1000 genes encoding such receptors in the human genome [2], and these receptors regulate virtually all known physiological processes in mammals. Their central importance and relevance to the current clinical practice of medicine is reflected in the plethora of drugs that target these receptors. A substantial proportion of all worldwide prescription drug sales today are generally attributed to drugs that target the 7TM receptors, either directly or indirectly (e.g. selective serotonin reuptake inhibitors or angiotensin-converting enzyme inhibitors) as agonists or antagonists [3].

Moreover, we know a great deal about the molecular properties of these receptors: how they signal, how their function is regulated, and how they adapt to changing physiological circumstances over short and chronic time frames. Most of this information has been discovered during the past 30 years. Yet, most of our current generation of students and fellows take this enormous body of information largely for granted and know little about how we came to our current understanding of this remarkable superfamily of receptor molecules. In fact, many of my trainees refer to literature references before ~2000 as 'old' references. It is in this context that I set out here to provide an admittedly very personal and brief account of discoveries and developments that occurred largely during the 1970s and 1980s and that served to establish many of the 'central dogmas' that now underpin the field. Although knowing the history of how current concepts emerged can deepen our appreciation and understanding of present ideas, perhaps even more importantly, it can help us to understand their limitations and even suggest avenues for future research.

Early antecedents

Although the earliest discussion of 'binding' of biologically active molecules to specific sites on cells is generally attributed to Ehrlich, it was Langley and his student Dale, working together during the first decade of the 20th century, who first explicitly stated the idea of a 'receptive substance' on reactive cells [4]. Their deductions were based on classical physiological and pharmacological experiments using isolated skeletal or smooth muscle preparations, or submandibular salivary glands, and combinations of adrenoceptor or acetylcholine receptor agonists and antagonists [5,6]. It is interesting to note that their early work, performed almost a century ago, involved the study of receptors that were representative of what we now know to be two of the most important families of receptors, the 7TM receptors (adrenoceptors and muscarinic acetylcholine receptors) and the ion channel receptors (nicotinic acetylcholine receptors).

Corresponding author: Robert J. Lefkowitz (lefko001@receptor-biol.duke.edu).
Available online 2 July 2004

 0165-6147/$ - see front matter © 2004 Elsevier Ltd. All rights reserved. doi:10.1016/j.tips.2004.06.006

During the next half century, from ~1920 to 1970, the concepts of Langley and Dale were developed into classical receptor theory by such giants of pharmacology as Clark, Ariens, Stephenson, Black and Furchgott. These scientists applied the law of mass action to classical dose–response data and developed the notions of affinity and efficacy (among many others) in drug action [7].

Biochemical effectors of receptor action

The 1960s and 1970s witnessed the beginning of a merger between biochemistry and pharmacology. This was led by biochemists such as Sutherland, Krebs and Rodbell, who turned their attention to studying the molecular basis of hormone and drug action. This led, in relatively quick succession, to a series of seminal discoveries that were to shape all further work in the field. Sutherland discovered the second messenger cAMP, which mediates the actions of dozens of receptors, and the enzyme adenylyl cyclase, which is responsible for the synthesis of cAMP [8], whereas Krebs discovered the cAMP-dependent protein kinase, which is the proximate effector of cAMP action [9]. In 1971, Rodbell proposed the existence of a guanine nucleotide regulatory protein that serves as a transducer between hormone receptors and adenylyl cyclase [10]. Gilman and colleagues subsequently demonstrated the existence of this protein [11], and later purified the protein and named it G_s [12]. Taken together, these discoveries provided the first chain of molecular events that convey signaling from the outside to the inside of the cell. Moreover, they provided biochemical assays and 'readouts' of cell stimulation that measured receptor activity more directly than classical physiological systems that had typically measured contraction of muscle cells, beating of the heart or glandular secretion. Although such assays brought the observed responses nearer to the initial point of contact of neurotransmitters or drugs with the cell, the receptors themselves remained illusive.

Molecular era of receptor research

The era of molecular research on the receptors (studies that assess receptor properties directly, rather than inferring them from downstream effector function) can reasonably be dated to ~1970. A remarkable fact about this period, difficult to imagine now, is that despite decades of physiological and biochemical work, the physical existence of receptors remained controversial. A sense of this skepticism is apparent in a statement made by Ahlquist. A prominent pharmacologist of the time, Ahlquist had developed the important concept of distinct α- and β-adrenoceptors in 1948. However, in 1973 he wrote:

'This would be true if I were so presumptuous as to believe that α and β receptors really did exist. There are those that think so and even propose to describe their intimate structure. To me they are an abstract concept conceived to explain observed responses of tissues produced by chemicals of various structure' [13].

As a biochemist, Sutherland was more comfortable with the notion of receptors as physicochemical entities, but even he was not yet ready to accord them an independent existence distinct from the enzyme adenylyl cyclase, which he had discovered. In 1967 he wrote:

'It seems likely that in most and perhaps all tissues the β receptor and adenylyl cyclase are the same. The results of many previous studies have pointed to this conclusion, and we feel that the studies with the perfused rat heart have added further to its possible validity' [14].

As is so often the case, the way forward required the development of several novel technologies. Most pertinent to our discussion are radioligand binding and affinity labeling techniques for the receptors, detergent solubilization, affinity chromatography purification and lipid reconstitution. Although progress on individual fronts was made for several receptors, only in a few cases were all these methods applied successfully to a single receptor. The model systems that emerged as most thoroughly investigated during this period and that ultimately led to breakthroughs in our understanding of the ion channels and 7TM receptors, respectively, were the nicotinic acetylcholine receptor [15], and the β_2-adrenoceptor [16,17] and rhodopsin [17,18].

Why did these systems emerge as the key prototypes? In the case of the nicotinic acetylcholine receptor, the availability of the electric organs of electric fish provided an extraordinarily rich source of these receptors [19]. In fact, purified membranes prepared from these organs contained the receptors at a specific activity that was almost comparable to that of pure receptor protein [19]. An analogous situation exists for rhodopsin, which constitutes ~90% of the protein in bovine retinal rod preparations [18]. Thus, in both of these cases, almost no purification was required after solubilization. Moreover, as discussed later, rhodopsin was not even viewed as a model for receptors until the cloning of the β_2-adrenoceptor revealed their homologous structures.

My reasons for studying the β_2-adrenoceptor in the early 1970s, and later the other adrenoceptors, was based on several considerations. First, the β_2-adrenoceptor was one of several receptors that activate adenylyl cyclase, which had been studied frequently in the past by luminaries such as Sutherland [14] and Krebs [9]. Second, I wanted to study a receptor with obvious clinical cardiovascular relevance. Third, the advent of β-blockers in clinical practice in the late 1960s, based on the prominent work of Black [20], had also helped focus my attention in this direction. Finally, there were some practical considerations. It seemed obvious that to probe the receptors it would be necessary to develop several novel technologies. The availability of dozens of β-adrenoceptor agonists and antagonists of varied structure provided an attractive collection of starting materials from which to develop radioligands, affinity labeling reagents and affinity chromatography matrices, among others.

Radioligand binding

If a single technical advance can be said to have opened the door to the molecular era of receptor research, it was the development of radioligand binding methods during the 1970s. These methods, which were developed at about the same time for several receptors now known to be 7TM receptors (recently reviewed for the opioid peptide receptors [21]) and for the nicotinic acetylcholine receptor, transformed the field of receptor research. These methods

led rapidly to new insights into the dynamic regulation of receptor number and properties, and insights into their molecular coupling properties. For example, these methods were used to develop approaches to quantitating and analyzing receptor coupling to G proteins. Thus, the 'ternary complex model' for understanding receptor interaction with G proteins [22] and the quantitation of high (G-protein-coupled) and low (not coupled) affinity states of the receptor [23] were applied widely and are used today (Figure 1). Ligand binding approaches also led to a rapid expansion of the roster of receptor subtypes and provided the first evidence for direct alterations in receptors during the process of desensitization (see later) [24].

Receptor purification and reconstitution

Perhaps the most important consequence of the ability to tag receptors directly was that it opened the way to receptor purification. In contrast to the situation with the nicotinic acetylcholine receptor and rhodopsin, no highly enriched sources of most GPCRs exist. However, working together with Joseph Pitha and Marc Caron, we developed a highly effective and specific affinity chromatography resin for the purification of the β_2-adrenoceptor [25–27]. During the next several years, postdoctoral students John Regan and Fredrik Leeb-Lundberg created specific affinity resins to purify both the α_{2A}-adrenoceptor [28] and α_{1B}-adrenoceptor [29] subtypes. By coupling these affinity chromatography procedures with more-conventional chromatographic procedures, we were able to purify first the β_2-adrenoceptor [25–27] and then the α_{2A}- [28] and α_{1B}-adrenoceptors [29] essentially to homogeneity (Figure 2). This required purification in excess of 100 000 fold. Consequently, only small amounts of protein were obtained, ranging from ~25 to 50 micrograms, often after pooling material from several runs.

The purified β_2-adrenoceptor consisted of a single polypeptide chain, which was variably glycosylated and phosphorylated, with a M_r of ~60 000 daltons [25–27]. β-Adrenoceptor ligands bound to the isolated receptor protein with appropriate specificity and stereospecificity. However, skepticism still persisted as to whether the isolated protein in fact constituted the true receptor. Such skepticism was finally put to rest only when the isolated receptor protein was successfully reconstituted in phospholipid vesicles and shown to be functionally active. In initial experiments, the receptor protein was shown to convey β-adrenoceptor responsiveness on the adenylyl cyclase of *Xenopus* laevis erythrocytes (which contain virtually no β-adrenoceptors) [30]. In subsequent experiments, the receptor was reconstituted successfully in vesicles with purified G_s and the catalytic moiety of adenylyl cyclase [31,32]. A fully functioning system was thereby established in which catecholamines could stimulate the enzyme. These experiments not only validated the functional identity of the isolated receptor protein but also defined the minimum number of components necessary to comprise a fully functional agonist-sensitive adenylyl cyclase system.

Receptor cloning

Although only relatively small amounts of β_2-adrenoceptor were obtained by purification, with advances in microsequencing techniques by the mid-1980s, this was sufficient to obtain small stretches of peptide sequence from peptides derived from the receptors. This enabled the design of oligonucleotide probes that could be used to clone the genes and/or cDNAs encoding the receptors. This led, in 1986, to an important breakthrough: the cloning of the gene and cDNA encoding the hamster β_2-adrenoceptor [33]. This was a difficult process because of the scarcity of cDNA clones for the receptors in cDNA libraries (only a few per million recombinants) and the fact that most of the libraries at that time were plasmid libraries, containing only a few hundred thousand independent recombinants.

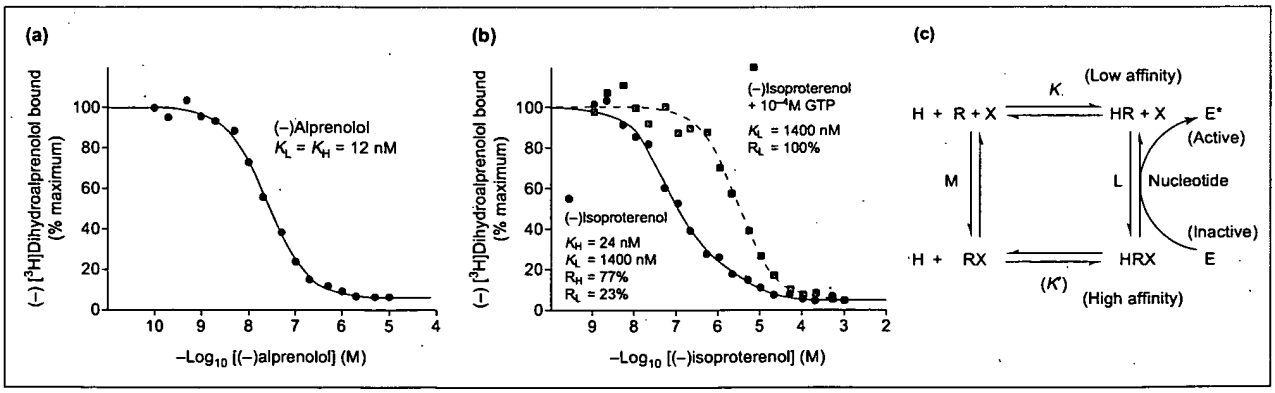


Figure 1. Computerized curve fitting of binding data from competition of the β-adrenoceptor radioligand [3H]dihydroalprenolol by **(a)** (–)alprenolol and **(b)** (–)isoproterenol (also known as isoprenaline) in the presence and absence of GTP. Reproduced, with permission, from [23]. These ligand-displacement curves illustrate striking differences obtained with an agonist (isoproterenol) and an antagonist (alprenolol). The experiments were performed using membranes from frog erythrocytes. The alprenolol curve is steep and uniphasic whereas the isoproterenol curve is biphasic and can be resolved into two components, one of high affinity and one of low affinity as indicated by K_H and K_L. R_H and R_L refer to the percentages of receptor sites of high affinity and low affinity, respectively, for agonist as analyzed by nonlinear least squares curve fitting. In the presence of GTP, the agonist curve shifts to the right and becomes steeper, and now all of the receptors are found in the low-affinity state. **(c)** The ternary complex model, where H is the hormone or agonist, R is the receptor, X is an additional component in the membrane (subsequently shown to be the guanine nucleotide regulatory protein) and E is an effector such as adenylyl cyclase. Modified from [22]. The classical ternary complex model indicates the initial formation of a low-affinity complex of agonist or hormone with receptor followed by a second step in which the agonist–receptor complex couples to a third component represented as X. A guanine nucleotide disrupts the high-affinity complex reverting it to the low-affinity form. X was predicted to be the guanine nucleotide regulatory protein. Thus, the model provides the conceptual framework for coupling the effects of guanine nucleotide on ligand–receptor interaction with the activation of the nucleotide regulatory protein and adenylyl cyclase.

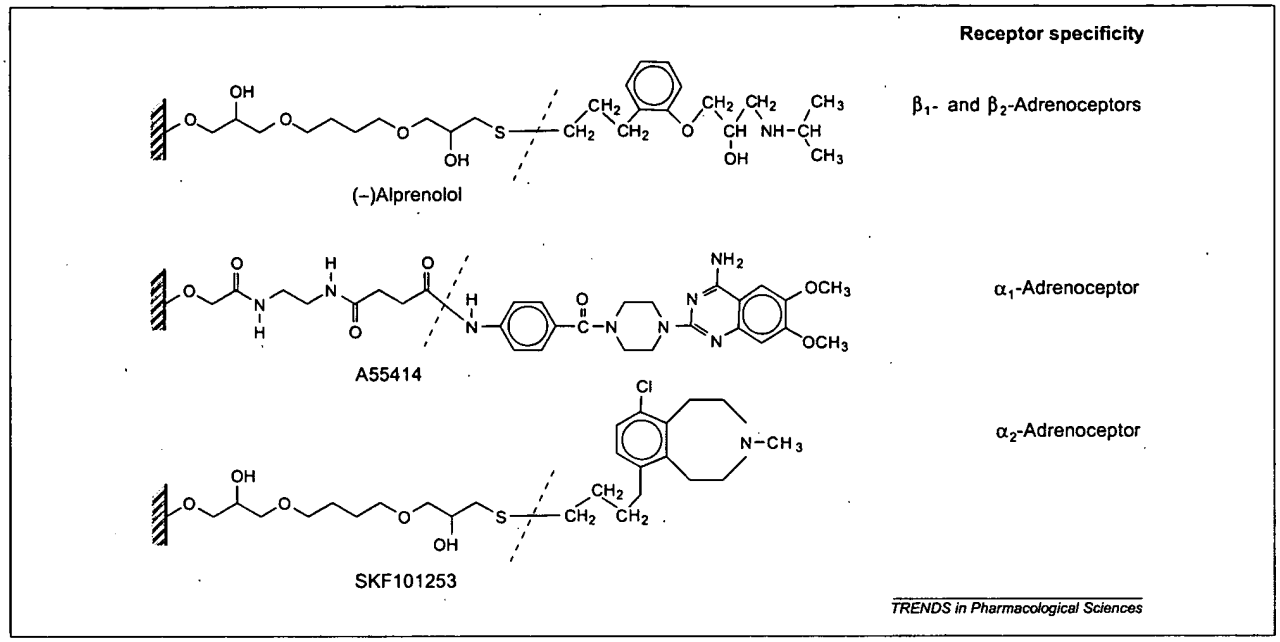


Figure 2. Biospecific affinity chromatography supports for the purification of adrenoceptors. The indicated β- or α-adrenoceptor antagonists were coupled to Sepharose beads and then used for affinity chromatography purification of the receptors. Solubilized receptor preparations were applied to the columns and the receptors adsorbed by high-specific biospecific interactions with the immobilized ligands. After appropriate washing procedures the purified receptors were eluted with appropriate α- or β-adrenoceptor ligands.

The key breakthrough came when Brian Kobilka decided almost in desperation to construct a genomic library for screening. This library was screened simultaneously by our laboratory and by our collaborator, Richard Dixon at Merck. Immediately both laboratories were successful in pulling out overlapping clones that encoded all the peptide sequences we had previously obtained. This led rapidly to several totally unexpected discoveries. The first was that the gene encoding the β_2-adrenoceptor was intronless [33,34]. Thus, the clones quickly yielded the entire deduced amino acid sequence of the receptor from a single exon. This feature of an intronless gene subsequently turned up again when we cloned the α_2-adrenoceptors [35], and is also a feature of several other members of the 7TM receptor gene family (e.g. muscarinic acetylcholine receptors [36] and dopamine D1 receptors [37], among many others). The second and more important unanticipated finding was that the β_2-adrenoceptor shared sequence homology and a predicted 7TM architecture with the visual pigment rhodopsin [33].

It is difficult today to appreciate how surprising this latter discovery was in 1986. Rhodopsin had been the subject of intense biochemical investigation for many years because of its centrality in the visual process and its remarkable abundance. This led to the elucidation of its entire amino acid sequence in 1982 by conventional protein sequencing [38,39]. By 1986, studies of the mechanisms of rhodopsin function had revealed its linkage to the G-protein transducin in a light-dependent fashion. The functional analogies with the β_2-adrenoceptor and G_s were already well appreciated at the time [18]. Nonetheless, a structural relationship between rhodopsin and the β_2-adrenoceptor, or other GPCRs, was not anticipated. Rhodopsin was not viewed as a model for other receptors. In fact, when the sequence of rhodopsin was determined, and the predicted 7TM arrangement found, it was immediately analogized to bacteriorhodopsin [38,39]. This bacterial protein is a light-driven proton pump. In fact, a leading hypothesis at this time was that all light-sensitive proteins might possess such a 7TM structure and that it might in fact represent the hallmark of such proteins [38–40].

As soon as we appreciated the structural similarities between the β_2-adrenoceptor and rhodopsin, we realized that it was likely that many or all GPCRs might share this arrangement [17,33,41]. This hypothesis was confirmed rapidly during the next several years, following our cloning of seven additional adrenoceptors [35,42–48]. Furthermore, Numa cloned the M_1 and M_2 muscarinic acetylcholine receptors [49] and Nakanishi cloned the NK_1 tachykinin receptor [50]. I never imagined, however, that the superfamily of 7TM receptors would grow so large and diverse. Members of the family are now known to encode receptors for ligands as diverse as odorants, tastants, Ca^{2+}, trace amines, lipids, peptides and even HIV [1]. Other than the adrenoceptors, few GPCRs have been purified before their cloning. Rather, their genes were cloned based on suspected or serendipitous homology with already-cloned members of the family, beginning with the adrenoceptors. Thus, the gene family expanded at an ever-accelerating pace as more and more receptors were added.

In many cases, the functional identity of these cloned receptors was not known, giving rise to the term 'orphan receptors'. The first example of this was the receptor clone 'G21', which we isolated from a size-selected human genomic DNA library shortly after the cloning of the β_2-adrenoceptor [51]. The clone was obtained by low-stringency hybridization with the β_2-adrenoceptor cDNA.

Because it represented the most strongly hybridizing signal in genomic DNA other than the β_2-adrenoceptor itself, we reasonably anticipated that it would encode the closely related β_1-adrenoceptor. In fact, expression studies indicated that it did not. A year later we demonstrated that the intronless gene encoded the 5-HT_{1A} receptor, which, in essence, was the first example of 'deorphanizing' such a receptor [52]. Today, a large majority of the thousand or so genes for 7TM receptors are 'orphans' and includes most of the olfactory receptors [53]. From our current perspective it is clear that the difficult and not very glamorous work of the 1970s and early 1980s, which ultimately led to the purification of the β_2-adrenoceptor and other adrenoceptors, indeed paid rich dividends. The cloning of the β_2-adrenoceptor and the discovery that it and rhodopsin were the founding members of a huge receptor gene family laid the foundation for all future work on this remarkable gene family.

Mutagenesis and chimeric receptors

With the availability of clones for a rapidly increasing number of receptors, the scientific community turned its attention to unraveling the structural features that determine receptor function. Extensive mutagenesis studies carried out in many laboratories quickly established the major principles [54,55]. The internal loops of the receptors, particularly those regions in closest proximity to the plasma membrane in addition to the proximal portions of the C-terminal tail, engaged the G proteins, whereas ligand binding was localized to residues in the membrane-spanning regions or the external loops. Interesting specializations were discovered. For example, within the thyrotropin-stimulating hormone (TSH) and gonadotrophin receptors, a markedly extended N-terminal peptide binds the large protein ligands [56]. In the case of the protease-activated receptors (PAR1–4), like the thrombin receptor [57], the proteolytic cleavage of the N-terminus of the receptor actually uncovers the peptide ligand at the new N-terminus.

In the early years of these investigations, we took the approach of creating 'chimeric' receptors as a means of probing receptor structure–function relationships. For example, having cloned the β_2-adrenoceptor, which stimulates adenylyl cyclase activity, and the α_{2A}-adrenoceptor, which inhibits adenylyl cyclase activity, we focused on identifying which regions of the molecules controlled which functions. A small segment of the β_2-adrenoceptor encompassing its third cytoplasmic loop, when spliced into the α_2-adrenoceptor, reversed the G-protein coupling and biological function of the α_2-adrenoceptor [58] (Figure 3). Thus, this chimeric $\alpha_2\beta_2$-adrenoceptor now activated rather than inhibited adenylyl cyclase but retained the pharmacological specificity (ligand-binding pattern) of the α_2-adrenoceptor. In this way, the regions of the molecules responsible for G-protein coupling (e.g. third cytoplasmic loop) and ligand binding (membrane-spanning domains) were delineated. With the further application of this approach, such functionally important structural features were subsequently defined with increasingly fine detail. Other approaches such as site-directed mutagenesis were also revealing [54,55].

Constitutively active mutant receptors

Another serendipitous consequence that followed the chimeric receptor work was the discovery of constitutively active mutant receptors. Thus, unexpectedly, when another fellow in my laboratory, Susanna Cotecchia, replaced just four residues in the third cytoplasmic loop of the α_{1B}-adrenoceptor with residues from the β_2-adrenoceptor, the α_{1B}-adrenoceptor acquired the ability to signal constitutively: that is, in the absence of agonist [59]. It was ultimately found that virtually any substitution at a particular site, specifically A293 in the α_{1B}-adrenoceptor, resulted, to some extent, in this activating effect [60]. This seemed to be due to the mutations abrogating important intramolecular interactions that normally keep the receptor constrained in an inactive conformation. The conformational effects of the mutagenesis presumably mimicked those normally induced by agonists. Subsequently, naturally occurring mutations in several receptors were found to lead to constitutive activation and to various diseases [61].

Moreover, studies of the constitutive activity of receptors led to a convincing demonstration, both *in vitro* [62] and *in vivo* [63], of the phenomenon of inverse agonism. This special property of some drugs, previously classified as classical or neutral antagonists, leads to a reversal of receptor constitutive activity [62,63], presumably by selectively binding and stabilizing an inactive state of the receptor. Inverse agonists might have interesting and as yet undiscovered therapeutic activities, because a generally low level of constitutive activity is a general feature of most receptors.

Arrestins and G-protein-coupled receptor kinases (GRKs): a universal mechanism of 7TM receptor regulation

As described earlier, the cloning of the β_2-adrenoceptor in 1986 and the appreciation that it and rhodopsin were structurally related, founding members of a large receptor gene family represented an important milestone. However, almost simultaneously two other distinct streams of research were converging that would also reveal the close relationship between these two receptors, and by obvious extension, a relationship that would ultimately extend to the large 7TM receptor family. This research concerned the discovery of a universal mechanism that regulates receptor function: stimulus-dependent receptor phosphorylation followed by arrestin binding. By the mid-1980s, it had become clear that both rhodopsin [64] and the β_2-adrenoceptor [65] were phosphorylated in a stimulus-dependent way, and that this phosphorylation seemed to be related to the process of receptor inactivation or desensitization. In the case of rhodopsin, the enzyme responsible for the phosphorylation was referred to as rhodopsin kinase [66]. In the case of the β_2-adrenoceptor, both the cAMP-dependent kinase (PKA) [67] and a novel cAMP-independent kinase [which we called the β-adrenoceptor kinase (βARK)] appeared to contribute to agonist-dependent phosphorylation of the receptor [68]. PKA appeared to mediate so-called 'heterologous' or agonist nonspecific desensitization [67], whereas βARK mediated agonist-specific or 'homologous' desensitization [68].

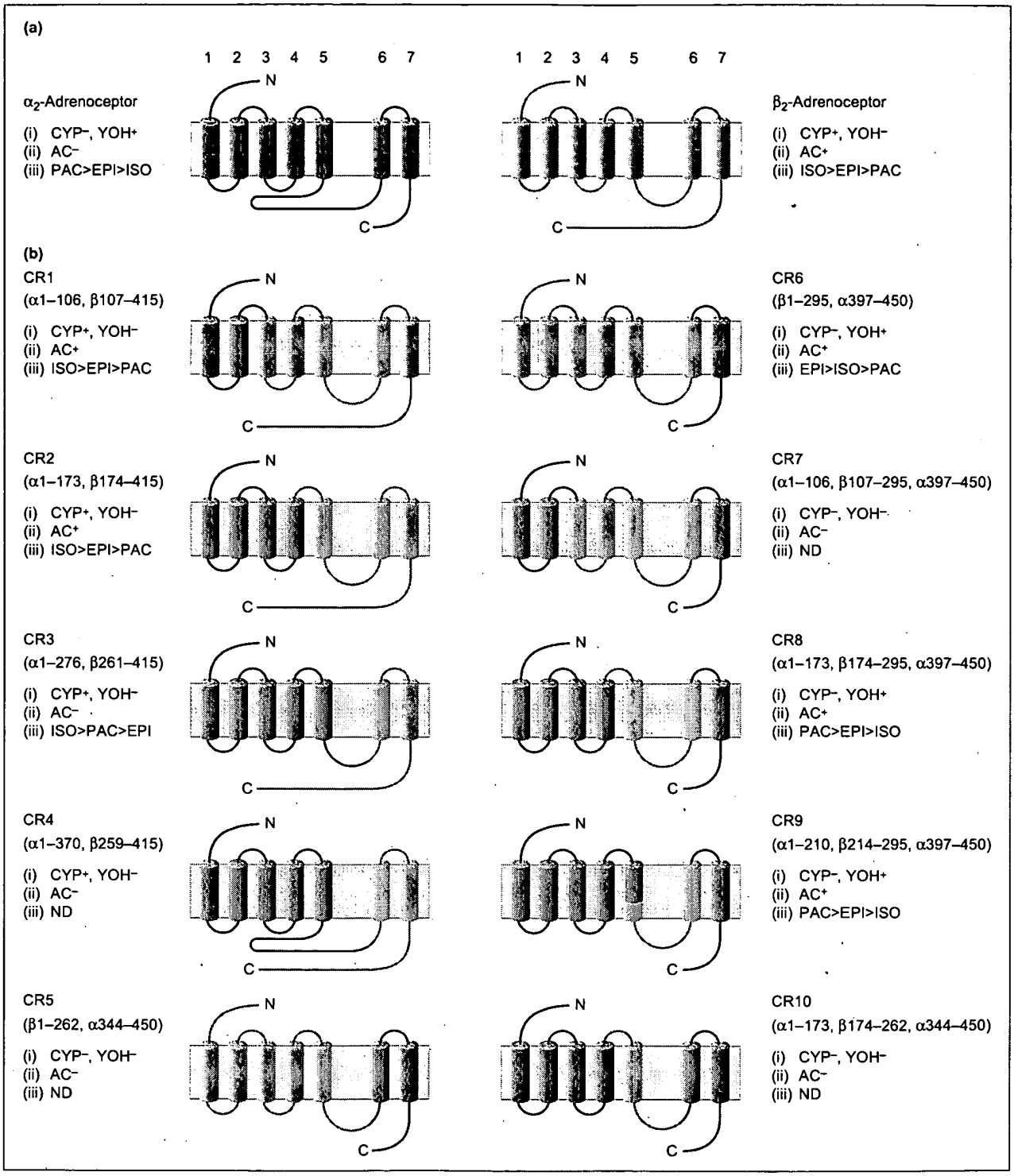


Figure 3. Chimeric $\alpha_2\beta_2$-adrenoceptors. **(a)** The wild-type α_2-adrenoceptor and the wild-type β_2-adrenoceptor. The hydrophobic domains are shown as forming α-helices that span the plasma membrane. These putative α-helices are numbered 1–7 from the N-terminus (extracellular) to the C-terminus (intracellular). **(b)** Chimeric receptors (CRs) formed from combinations of the wild-type α_2-adrenoceptor (blue) and β_2-adrenoceptor (green). The α_2-adrenoceptor (α) and β_2-adrenoceptor (β) amino acid sequences from the N- to C-terminus are indicated in parentheses beside each chimeric receptor. Beside each receptor is a summary of the functional characteristics of the receptor expressed in *Xenopus laevis* oocytes or COS-7 cells (or both). The functional properties include: (i) the ability to bind the β_2-adrenoceptor antagonist ^{125}I-labeled cyanopindolol (CYP) and the α_2-adrenoceptor-specific antagonist [^{3}H]yohimbine (YOH); (ii) the ability to couple to G_s and activate adenylyl cyclase (AC) after stimulation by epinephrine (also known as adrenaline); and (iii) the relative potency of the α_2-and β_2-adrenoceptor agonist epinephrine (EPI), the α_2-adrenoceptor-specific agonist *p*-aminoclonidine (PAC), and the β-adrenoceptor-specific agonist isoproterenol [ISO (also known as isoprenaline)] for the receptor as determined by ligand binding studies or adenylyl cyclase activation (or both). Chimeric receptor genes were constructed by splicing desired restriction endonuclease fragments from the wild-type receptor genes with synthetic oligonucleotide adapters. The most informative chimeras are CR8 and CR9, which activate adenylyl cyclase despite containing only a small insert

In 1986, Herman Kuhn and collaborators reported that an abundant retinal protein, previously known as S antigen or '48K protein', bound to phosphorylated rhodopsin, thereby interfering with its subsequent coupling to transducin [69]. The protein was renamed arrestin by Liebman [70]. At the time, Jeff Benovic, a graduate student in my laboratory, was working on purification of βARK from bovine brain [71]. He had devised an assay in which enzyme preparations were used to phosphorylate purified (non-recombinant) mammalian β_2-adrenoceptor reconstituted in phospholipid vesicles. In addition to ^{32}P incorporation, we followed the loss of the ability of the receptor to convey catecholamine stimulation to co-reconstituted G_s (assessed as GTPase activity), an *in vitro* model of desensitization. To our initial surprise and disappointment, the more purified the enzyme preparations became, the less they were able to 'desensitize' the receptor [72].

We had already been toying with the idea that perhaps a cofactor that was necessary for the receptor inactivation was being lost during the purification process. Immediately upon reading the Kuhn report [69], we hypothesized that a factor analogous to retinal arrestin might be the answer. Contacting him, we obtained some of his retinal arrestin and demonstrated that it restored the desensitizing ability of the βARK preparations, albeit at much higher concentrations than were necessary in analogous reconstituted rhodopsin systems [72].

Emboldened by these results, we sought the hypothetical related protein that we named β-arrestin. Our attempts at biochemical isolation failed. However, within several years, Shinohara cloned retinal arrestin [73]. Using an arrestin clone, in low-stringency screens, we isolated clones for the putative β-arrestin1 [74], and subsequently for β-arrestin2 [75]. β-Arrestin1 and 2 are also referred to as arrestin2 and 3, whereas visual arrestin is also referred to as arrestin1. In contrast to the restricted expression of retinal arrestin, β-arrestins1 and 2 are expressed ubiquitously [75]. In reconstituted systems, the β-arrestins were orders of magnitude more potent than visual arrestin in desensitizing the β_2-adrenoceptor-stimulated G_s and vice versa for retinal arrestin in reconstituted rhodopsin systems [75] (Figure 4).

Contemporaneously, we demonstrated that purified preparations of βARK could catalyze light-dependent phosphorylation of rhodopsin and that rhodopsin kinase could phosphorylate the β_2-adrenoceptor in an agonist-dependent fashion [76]. Both kinases, however, demonstrated marked preference for their physiological receptor substrates. Moreover, the cloning of the β_2-adrenoceptor had suggested that, as with rhodopsin, the relevant sites of stimulus-dependent phosphorylation were probably localized to the serine- and threonine-rich carboxyl terminus [33,76]. All these findings suggested to us that, as with the receptors themselves, a gene family of GRKs might exist, a speculation confirmed by our subsequent cloning of βARK [77] and rhodopsin kinase [78]. The two enzymes defined a new subfamily of serine/threonine kinases [79]. Subsequently, we, and others, cloned additional members of the family, which is now known to contain seven kinases, referred to as GRKs1–7 [80–82]. GRK1 (rhodopsin kinase) and GRK7 are expressed only in retinal rods and cones, respectively. GRKs2 (βARK), 3, 5 and 6 are expressed ubiquitously, whereas GRK4 has a much more restricted pattern of expression.

Today, it is appreciated that, almost without exception, 7TM receptors are regulated by GRKs and arrestins. That only a handful of enzymes and only two forms of β-arrestin can do this emphasizes the remarkable ability of these proteins to interact with literally hundreds of different receptors. The key to their physiological specificity, of course, is that they interact only with the activated forms of the receptors. Details of which kinases phosphorylate which receptors in any given tissue, however, remain largely unknown.

β-Arrestins and GRKs in endocytosis and signaling

Completely unsuspected at the time of their discovery was the fact that the biological functions of β-arrestins and GRKs extend well beyond 'desensitization' of the receptors. Thus, in recent years, β-arrestins have been shown to serve as agonist-dependent adaptors that link the receptors both to elements of the clathrin-dependent endocytic machinery and to a growing list of signaling proteins [83–85]. Interaction of β-arrestins with clathrin, adaptor protein 2 (AP2), the small G protein ADP-ribosylation factor 6 (ARF6) and its exchange factor ARF nucleotide binding site opener (ARNO), and N-ethylmaleimide-sensitive fusion protein (NSF) have all been implicated in this endocytic function [83–85]. Moreover, the process appears to be regulated by β-arrestin ubiquitination, mediated by the E3 ubiquitin ligase MDM2 [86].

The list of 7TM receptor-stimulated signaling pathways in which β-arrestins have been implicated includes non-receptor tyrosine kinases of the Src family [87], mitogen-activated protein (MAP) kinases, including extracellular signal-regulated kinases 1,2, (ERK1,2) [87,88], p38 [89] and Jun N-terminal kinase 3 (JNK3) [90], phosphatidylinositol 3-kinase [91] and protein kinase B [91]. β-Arrestins have also been found to play analogous roles for the receptor tyrosine kinase IGF1 receptor [92,93]. The most thoroughly studied systems have been the ERKs. Here, β-arrestin2 appears to function as a scaffold that binds not only the MAP kinase but also upstream kinases in the cascade [87,88]. Not only does the β-arrestin facilitate activation of the MAP kinase but it also directs the activated enzyme to cytosolic locations (e.g. endocytic vesicles), thereby excluding it from the nucleus [88,94,95]. MAP kinases activated in this way probably mediate cellular events distinct from, and occurring more rapidly than, those associated with the classical transcriptional regulatory effects of MAP kinases. For example, β-arrestin-mediated MAP kinase activation appears to play a role in some forms of chemotaxis [96].

from the β_2-adrenoceptor. The activity of CR8 and CR9, which bind ligands with an α_2-adrenoceptor specificity, demonstrates their ability to activate adenylyl cyclase, like the β_2-adrenoceptor, and implicates the third cytoplasmic loop of the receptor as being a key determinant of G-protein-coupling specificity. Abbreviation: ND, not determined. Redrawn, with permission, from [58].

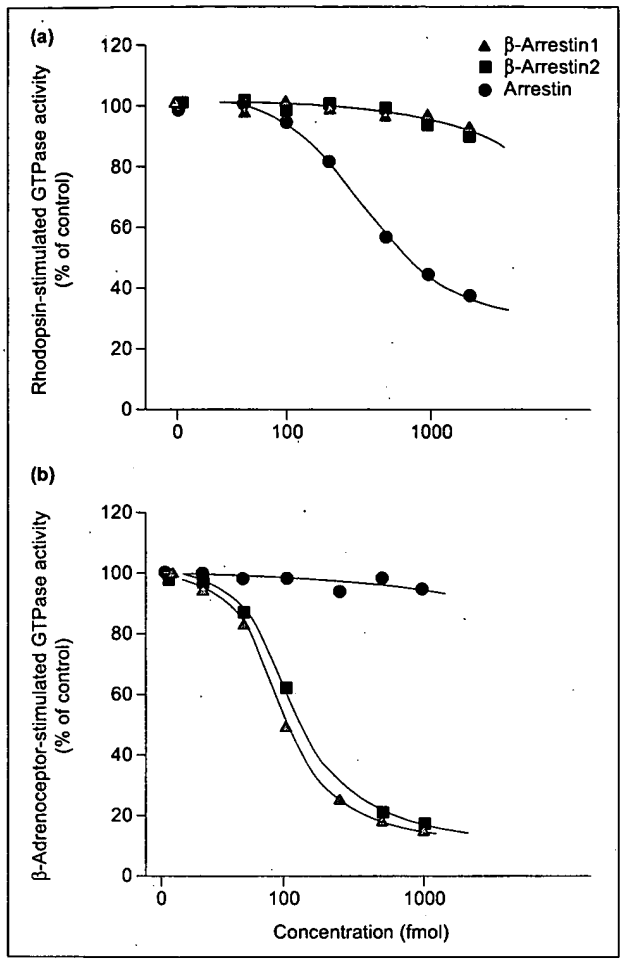


Figure 4. Inhibition of β_2-adrenoceptor function and rhodopsin function by β-arrestin1, β-arrestin2 and arrestin. See [75] for experimental details. Phosphorylated rhodopsin- (a) or β_2-adrenoceptor- (b) stimulated GTPase in transducin or G_s, respectively, was inhibited by recombinant versions of the various members of the arrestin family. These data demonstrate marked specificity of the β-arrestins for phosphorylated β_2-adrenoceptors and of arrestin for phosphorylated rhodopsin. Reproduced, with permission, from [75].

Future perspectives and opportunities

The discoveries of the past 30 years have transformed receptors from abstract physiological concepts into physicochemical entities. They have revealed pervasive, even universal, principles concerning their structure, function and regulation. Today, there are many hot topics in the field of 7TM receptor biology, several of which are likely to lead to new and/or improved therapeutics. A few that I consider to be of most importance are listed below.

(i) Attempts to crystallize the receptors and associated signaling proteins and determine their three-dimensional structure. This will provide a molecular basis for designing many new therapeutic agents [97–99].

(ii) Homo- and heterodimerization of the receptors. Highly controversial just a few years ago, the idea now seems widely accepted but the physiological significance is still unclear [100].

(iii) 'Deorphanizing' the dozens of orphan receptors found in the sequence of the human genome. This should ultimately yield attractive new drug targets [101].

(iv) Discovery of novel signaling mechanisms: for example, β-arrestins as signal transducers [83–90] or other signaling proteins that interact directly with the receptors [102]. The realization that G proteins represent only one mechanism of 7TM receptor signaling opens the possibility of developing drugs that are specific for newly discovered pathways, with potentially novel therapeutic actions.

(v) Pharmacogenomics. Polymorphisms in several 7TM receptors have already been described with important therapeutic and clinical implications. Such polymorphisms can modify the risk or course of a disease and determine the therapeutic response to particular drugs. For example, patients with a single amino acid change in the chemokine CCR5 receptor, which functions as a co-receptor for HIV, are resistant to infection of their T cells [103]. Patients with heart failure respond better to β-adrenoceptor antagonists if they have arginine rather than glycine at position 389 of the β_1-adrenoceptor [104]. Given the huge number of drugs that target 7TM receptors this area is likely to be especially fertile.

As exciting and surprising as the past several decades of research in this area have been, it seems likely that the years ahead might hold even more interesting and unexpected findings.

References

1 Pierce, K.L. *et al.* (2002) Seven-transmembrane receptors. *Nat. Rev. Mol. Cell Biol.* 3, 639–650
2 Fredriksson, R. *et al.* (2003) The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol. Pharmacol.* 63, 1256–1272
3 Gudermann, T. *et al.* (1995) Receptors and G proteins as primary components of transmembrane signal transduction. Part 1. G-protein-coupled receptors: structure and function. *J. Mol. Med.* 73, 51–63
4 Limbird, L.E. (1996) Historical perspective. In *Cell Surface Receptors: A Short Course on Theory and Methods* (Limbird, L.E., ed.), pp. 1–26, Kluwer Academic
5 Langley, J.N. (1901) Observations on the physiological action of extracts of the supra-renal bodies. *J. Physiol. (Lond.)* 17, 231–256
6 Dale, H.H. (1906) On some physiological actions of ergot. *J. Physiol. (Lond.)* 34, 163–206
7 Furchgott, R.F. (1964) Receptor mechanisms. *Annu. Rev. Pharmacol.* 4, 21–50
8 Sutherland, E.W. (1971) Cyclic AMP. In *An Introduction* (Robison, G.A. *et al.*, eds), pp. 5–13, Academic Press
9 Walsh, D.A. *et al.* (1968) An adenosine 3',5'-monophosphate-dependent protein kinase from rabbit skeletal muscle. *J. Biol. Chem.* 243, 3763–3765
10 Rodbell, M. *et al.* (1971) The glucagon-sensitive adenyl cyclase system in plasma membranes of rat liver. V. An obligatory role of guanylnucleotides in glucagon action. *J. Biol. Chem.* 246, 1877–1882
11 Ross, E.M. and Gilman, A.G. (1977) Resolution of some components of adenylate cyclase necessary for catalytic activity. *J. Biol. Chem.* 252, 6966–6969
12 Gilman, A.G. (1987) G proteins: transducers of receptor-generated signals. *Annu. Rev. Biochem.* 56, 615–649
13 Ahlquist, R.P. (1973) Adrenergic receptors: a personal and practical view. *Perspect. Biol. Med.* 17, 119–122
14 Robison, G.A. *et al.* (1967) Adenyl cyclase as an adrenergic receptor. *Ann. New York Acad. Sci.* 139, 703–723
15 Devillers-Thiery, A. *et al.* (1993) Functional architecture of the nicotinic acetylcholine receptor: a prototype of ligand-gated ion channels. *J. Membr. Biol.* 136, 97–112
16 Dohlman, H.G. *et al.* (1991) Model systems for the study of seven-transmembrane-segment receptors. *Annu. Rev. Biochem.* 60, 653–688
17 Lefkowitz, R.J. *et al.* (1986) β-Adrenergic receptors and rhodopsin: shedding new light on an old subject. *Trends Pharmacol. Sci.* 7, 444–448

18 Stryer, L. (1986) Cyclic GMP cascade of vision. *Annu. Rev. Neurosci.* 9, 87–119
19 Sobel, A. *et al.* (1977) Large-scale purification of the acetylcholine-receptor protein in its membrane-bound and detergent-extracted forms from *Torpedo marmorata* electric organ. *Eur. J. Biochem.* 80, 215–224
20 Black, J. (1989) Nobel lecture in physiology or medicine–1988. Drugs from emasculated hormones: the principle of syntopic antagonism. *In Vitro Cell. Dev. Biol.* 25, 311–320
21 Snyder, S.H. and Pasternak, G.W. (2003) Historical review: Opioid receptors. *Trends Pharmacol. Sci.* 24, 198–205
22 De Lean, A. *et al.* (1980) A ternary complex model explains the agonist-specific binding properties of the adenylate cyclase-coupled beta-adrenergic receptor. *J. Biol. Chem.* 255, 7108–7117
23 Kent, R.S. *et al.* (1980) A quantitative analysis of beta-adrenergic receptor interactions: resolution of high and low affinity states of the receptor by computer modeling of ligand binding data. *Mol. Pharmacol.* 17, 14–23
24 Mickey, J. *et al.* (1975) Subsensitivity of adenylate cyclase and decreased beta-adrenergic receptor binding after chronic exposure to (-)isoproterenol *in vitro*. *J. Biol. Chem.* 250, 5727–5729
25 Caron, M.G. *et al.* (1979) Affinity chromatography of the beta-adrenergic receptor. *J. Biol. Chem.* 254, 2923–2927
26 Shorr, R.G. *et al.* (1981) Purification of the beta-adrenergic receptor. Identification of the hormone binding subunit. *J. Biol. Chem.* 256, 5820–5826
27 Benovic, J.L. *et al.* (1984) The mammalian beta 2-adrenergic receptor: purification and characterization. *Biochemistry* 23, 4510–4518
28 Regan, J.W. *et al.* (1986) Purification and characterization of the human platelet alpha 2-adrenergic receptor. *J. Biol. Chem.* 261, 3894–3900
29 Lomasney, J.W. *et al.* (1986) Mammalian alpha 1-adrenergic receptor. Purification and characterization of the native receptor ligand binding subunit. *J. Biol. Chem.* 261, 7710–7716
30 Cerione, R.A. *et al.* (1983) Reconstitution of beta-adrenergic receptors in lipid vesicles: affinity chromatography-purified receptors confer catecholamine responsiveness on a heterologous adenylate cyclase system. *Proc. Natl. Acad. Sci. U. S. A.* 80, 4899–4903
31 Cerione, R.A. *et al.* (1984) Reconstitution of a hormone-sensitive adenylate cyclase system. The pure beta-adrenergic receptor and guanine nucleotide regulatory protein confer hormone responsiveness on the resolved catalytic unit. *J. Biol. Chem.* 259, 9979–9982
32 May, D.C. *et al.* (1985) Reconstitution of catecholamine-stimulated adenylate cyclase activity using three purified proteins. *J. Biol. Chem.* 260, 15829–15833
33 Dixon, R.A. *et al.* (1986) Cloning of the gene and cDNA for mammalian beta-adrenergic receptor and homology with rhodopsin. *Nature* 321, 75–79
34 Kobilka, B.K. *et al.* (1987) Delineation of the intronless nature of the genes for the human and hamster beta 2-adrenergic receptor and their putative promoter regions. *J. Biol. Chem.* 262, 7321–7327
35 Kobilka, B.K. *et al.* (1987) Cloning, sequencing, and expression of the gene coding for the human platelet alpha 2-adrenergic receptor. *Science* 238, 650–656
36 Pepitoni, S. *et al.* (1997) Structure of the m1 muscarinic acetylcholine receptor gene and its promoter. *J. Biol. Chem.* 272, 17112–17117
37 Sunahara, R.K. *et al.* (1990) Human dopamine D1 receptor encoded by an intronless gene on chromosome 5. *Nature* 347, 80–83
38 Ovchinnikov, Y.A. (1982) Rhodopsin and bacteriorhodopsin: structure-function relationships. *FEBS Lett.* 148, 179–191
39 Hargrave, P.A. *et al.* (1983) The structure of bovine rhodopsin. *Biophys. Struct. Mech.* 9, 235–244
40 Henderson, R. and Unwin, P.N. (1975) Three-dimensional model of purple membrane obtained by electron microscopy. *Nature* 257, 28–32
41 Dohlman, H.G. *et al.* (1987) A family of receptors coupled to guanine nucleotide regulatory proteins. *Biochemistry* 26, 2657–2664
42 Frielle, T. *et al.* (1987) Cloning of the cDNA for the human beta 1-adrenergic receptor. *Proc. Natl. Acad. Sci. U. S. A.* 84, 7920–7924
43 Regan, J.W. *et al.* (1988) Cloning and expression of a human kidney cDNA for an alpha 2-adrenergic receptor subtype. *Proc. Natl. Acad. Sci. U. S. A.* 85, 6301–6305
44 Cotecchia, S. *et al.* (1988) Molecular cloning and expression of the cDNA for the hamster alpha 1-adrenergic receptor. *Proc. Natl. Acad. Sci. U. S. A.* 85, 7159–7163
45 Schwinn, D.A. *et al.* (1990) Molecular cloning and expression of the cDNA for a novel alpha 1-adrenergic receptor subtype. *J. Biol. Chem.* 265, 8183–8189
46 Lomasney, J.W. *et al.* (1990) Expansion of the alpha 2-adrenergic receptor family: cloning and characterization of a human alpha 2-adrenergic receptor subtype, the gene for which is located on chromosome 2. *Proc. Natl. Acad. Sci. U. S. A.* 87, 5094–5098
47 Lomasney, J.W. *et al.* (1991) Molecular cloning and expression of the cDNA for the alpha 1A-adrenergic receptor. *J. Biol. Chem.* 266, 6365–6369
48 Lefkowitz, R.J. and Caron, M.G. (1989) The Adrenergic Receptors. In *Proceedings of The Robert A. Welch Foundation Conferences on Chemical Research, XXXIII*, pp. 183–199
49 Fukuda, K. *et al.* (1987) Molecular distinction between muscarinic acetylcholine receptor subtypes. *Nature* 327, 623–625
50 Masu, Y. *et al.* (1987) cDNA cloning of bovine substance-K receptor through oocyte expression system. *Nature* 329, 836–838
51 Kobilka, B.K. *et al.* (1987) An intronless gene encoding a potential member of the family of receptors coupled to guanine nucleotide regulatory proteins. *Nature* 329, 75–79
52 Fargin, A. *et al.* (1988) The genomic clone G-21 which resembles a beta-adrenergic receptor sequence encodes the 5-HT1A receptor. *Nature* 335, 358–360
53 Buck, L. and Axel, R. (1991) A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell* 65, 175–187
54 Ostrowski, J. *et al.* (1992) Mutagenesis of the beta 2-adrenergic receptor: how structure elucidates function. *Annu. Rev. Pharmacol. Toxicol.* 32, 167–183
55 Strader, C.D. *et al.* (1994) Structure and function of G protein-coupled receptors. *Annu. Rev. Biochem.* 63, 101–132
56 Parmentier, M. *et al.* (1989) Molecular cloning of the thyrotropin receptor. *Science* 246, 1620–1622
57 Vu, T.K. *et al.* (1991) Molecular cloning of a functional thrombin receptor reveals a novel proteolytic mechanism of receptor activation. *Cell* 64, 1057–1068
58 Kobilka, B.K. *et al.* (1988) Chimeric alpha 2-,beta 2-adrenergic receptors: delineation of domains involved in effector coupling and ligand binding specificity. *Science* 240, 1310–1316
59 Cotecchia, S. *et al.* (1992) Discrete amino acid sequences of the alpha 1-adrenergic receptor determine the selectivity of coupling to phosphatidylinositol hydrolysis. *J. Biol. Chem.* 267, 1633–1639
60 Kjelsberg, M.A. *et al.* (1992) Constitutive activation of the alpha 1B-adrenergic receptor by all amino acid substitutions at a single site. Evidence for a region which constrains receptor activation. *J. Biol. Chem.* 267, 1430–1433
61 Spiegel, A.M. (1998) *G Proteins, Receptors, and Disease*, Humana Press
62 Costa, T. and Herz, A. (1989) Antagonists with negative intrinsic activity at delta opioid receptors coupled to GTP-binding proteins. *Proc. Natl. Acad. Sci. U. S. A.* 86, 7321–7325
63 Bond, R.A. *et al.* (1995) Physiological effects of inverse agonists in transgenic mice with myocardial overexpression of the beta 2-adrenoceptor. *Nature* 374, 272–276
64 Wilden, U. and Kuhn, H. (1982) Light-dependent phosphorylation of rhodopsin: number of phosphorylation sites. *Biochemistry* 21, 3014–3022
65 Stadel, J.M. *et al.* (1983) Catecholamine-induced desensitization of turkey erythrocyte adenylate cyclase is associated with phosphorylation of the beta-adrenergic receptor. *Proc. Natl. Acad. Sci. U. S. A.* 80, 3173–3177
66 Shichi, H. and Somers, R.L. (1978) Light-dependent phosphorylation of rhodopsin. Purification and properties of rhodopsin kinase. *J. Biol. Chem.* 253, 7040–7046
67 Benovic, J.L. *et al.* (1985) Phosphorylation of the mammalian beta-adrenergic receptor by cyclic AMP-dependent protein kinase. Regulation of the rate of receptor phosphorylation and dephosphorylation by agonist occupancy and effects on coupling of the receptor to the stimulatory guanine nucleotide regulatory protein. *J. Biol. Chem.* 260, 7094–7101
68 Benovic, J.L. *et al.* (1986) Beta-adrenergic receptor kinase: identification of a novel protein kinase that phosphorylates the agonist-occupied form of the receptor. *Proc. Natl. Acad. Sci. U. S. A.* 83, 2797–2801
69 Wilden, U. *et al.* (1986) Phosphodiesterase activation by photoexcited rhodopsin is quenched when rhodopsin is phosphorylated and binds

the intrinsic 48-kDa protein of rod outer segments. *Proc. Natl. Acad. Sci. U. S. A.* 83, 1174–1178
70 Zuckerman, R. *et al.* (1985) Arrestin mediates ATP/ADP exchange and quench of cGMP phosphodiesterase activation. *Invest. Ophthalmol. Vis. Sci. (Suppl.)* 26, 45
71 Benovic, J.L. *et al.* (1987) Purification and characterization of the beta-adrenergic receptor kinase. *J. Biol. Chem.* 262, 9026–9032
72 Benovic, J.L. *et al.* (1987) Functional desensitization of the isolated beta-adrenergic receptor by the beta-adrenergic receptor kinase: potential role of an analog of the retinal protein arrestin (48-kDa protein). *Proc. Natl. Acad. Sci. U. S. A.* 84, 8879–8882
73 Shinohara, T. *et al.* (1987) Primary and secondary structure of bovine retinal S antigen (48-kDa protein). *Proc. Natl. Acad. Sci. U. S. A.* 84, 6975–6979
74 Lohse, M.J. *et al.* (1990) beta-Arrestin: a protein that regulates beta-adrenergic receptor function. *Science* 248, 1547–1550
75 Attramadal, H. *et al.* (1992) Beta-arrestin2, a novel member of the arrestin/beta-arrestin gene family. *J. Biol. Chem.* 267, 17882–17890
76 Benovic, J.L. *et al.* (1986) Light-dependent phosphorylation of rhodopsin by beta-adrenergic receptor kinase. *Nature* 321, 869–872
77 Benovic, J.L. *et al.* (1989) Beta-adrenergic receptor kinase: primary structure delineates a multigene family. *Science* 246, 235–240
78 Lorenz, W. *et al.* (1991) The receptor kinase family: primary structure of rhodopsin kinase reveals similarities to the beta-adrenergic receptor kinase. *Proc. Natl. Acad. Sci. U. S. A.* 88, 8715–8719
79 Lefkowitz, R.J. (1993) G protein-coupled receptor kinases. *Cell* 74, 409–412
80 Pitcher, J.A. *et al.* (1998) G protein-coupled receptor kinases. *Annu. Rev. Biochem.* 67, 653–692
81 Willets, J.M. *et al.* (2003) Non-visual GRKs: are we seeing the whole picture? *Trends Pharmacol. Sci.* 24, 626–633
82 Kohout, T.A. and Lefkowitz, R.J. (2003) Regulation of G protein-coupled receptor kinases and arrestins during receptor desensitization. *Mol. Pharmacol.* 63, 9–18
83 Shenoy, S.K. and Lefkowitz, R.J. (2003) Multifaceted roles of beta-arrestins in the regulation of seven-membrane-spanning receptor trafficking and signalling. *Biochem. J.* 375, 503–515
84 Perry, S.J. and Lefkowitz, R.J. (2002) Arresting developments in heptahelical receptor signaling and regulation. *Trends Cell Biol.* 12, 130–138
85 Ferguson, S.S. (2001) Evolving concepts in G protein-coupled receptor endocytosis: the role in receptor desensitization and signaling. *Pharmacol. Rev.* 53, 1–24
86 Shenoy, S.K. *et al.* (2001) Regulation of receptor fate by ubiquitination of activated beta 2-adrenergic receptor and beta-arrestin. *Science* 294, 1307–1313
87 Luttrell, L.M. *et al.* (1999) β-arrestin-dependent formation of β2-adrenergic receptor/Src kinase complexes and mitogenic signaling. *Science* 283, 655–661
88 Luttrell, L.M. *et al.* (2001) Activation and targeting of extracellular signal-regulated kinases by beta-arrestin scaffolds. *Proc. Natl. Acad. Sci. U. S. A.* 98, 2449–2454
89 Sun, Y. *et al.* (2002) Beta-arrestin2 is critically involved in CXCR4-mediated chemotaxis, and this is mediated by its enhancement of p38 MAPK activation. *J. Biol. Chem.* 277, 49212–49219
90 McDonald, P.H. *et al.* (2000) Beta-arrestin 2: a receptor-regulated MAPK scaffold for the activation of JNK3. *Science* 290, 1574–1577
91 Goel, R. *et al.* (2002) alpha-Thrombin induces rapid and sustained Akt phosphorylation by beta-arrestin1-dependent and -independent mechanisms, and only the sustained Akt phosphorylation is essential for G1 phase progression. *J. Biol. Chem.* 277, 18640–18648
92 Povsic, T.J. *et al.* (2003) Beta-arrestin1 mediates insulin-like growth factor 1 (IGF-1) activation of phosphatidylinositol 3-kinase (PI3K) and anti-apoptosis. *J. Biol. Chem.* 278, 51334–51339
93 Lin, F.T. *et al.* (1998) beta-arrestins regulate mitogenic signaling and clathrin-mediated endocytosis of the insulin-like growth factor I receptor. *J. Biol. Chem.* 273, 31640–31643
94 Tohgo, A. *et al.* (2002) beta-Arrestin scaffolding of the ERK cascade enhances cytosolic ERK activity but inhibits ERK-mediated transcription following angiotensin AT1a receptor stimulation. *J. Biol. Chem.* 277, 9429–9436
95 Tohgo, A. *et al.* (2003) The stability of the G protein-coupled receptor-beta-arrestin interaction determines the mechanism and functional consequence of ERK activation. *J. Biol. Chem.* 278, 6258–6267
96 Fong, A.M. *et al.* (2002) Defective lymphocyte chemotaxis in beta-arrestin2- and GRK6-deficient mice. *Proc. Natl. Acad. Sci. U. S. A.* 99, 7478–7483
97 Palczewski, K. *et al.* (2000) Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* 289, 739–745
98 Hirsch, J.A. *et al.* (1999) The 2.8 A crystal structure of visual arrestin: a model for arrestin's regulation. *Cell* 97, 257–269
99 Lodowski, D.T. *et al.* (2003) Keeping G proteins at bay: a complex between G protein-coupled receptor kinase 2 and Gbetagamma. *Science* 300, 1256–1262
100 Angers, S. *et al.* (2002) Dimerization: an emerging concept for G protein-coupled receptor ontogeny and function. *Annu. Rev. Pharmacol. Toxicol.* 42, 409–435
101 Lee, D.K. *et al.* (2003) Continued discovery of ligands for G protein-coupled receptors. *Life Sci.* 74, 293–297
102 Brady, A.E. and Limbird, L.E. (2002) G protein-coupled receptor interacting proteins: emerging roles in localization and signal transduction. *Cell. Signal.* 14, 297–309
103 Samson, M. *et al.* (1996) Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature* 382, 722–725
104 Mialet Perez, J. *et al.* (2003) Beta 1-adrenergic receptor polymorphisms confer differential function and predisposition to heart failure. *Nat. Med.* 9, 1300–1305

Articles of interest in other *Trends and Drug Discovery Today* journals

Advances in protein kinase B signalling: *AKT*ion on multiple fronts
Derek P. Brazil, Zhong-Zhou Yang and Brian A. Hemmings, *Trends in Biochemical Sciences* 29, 233–242

Targeted gene therapy: frontiers in the development of 'smart drugs'
Ingo H. Tarner, Ulf Müller-Ladner and C. Garrison Fathman, *Trends in Biotechnology* 22, 304–310

The enigmatic TRPCs: multifunctional cation channels
James W. Putney, Jr, *Trends in Cell Biology* 14, 282–286

Immune and nervous system CXCL12 and CXCR4: parallel roles in patterning and plasticity
robyn S. Klein and Joshua B. Rubin, *Trends in Immunology* 25, 306–314

Leukotrienes and atherosclerosis: new roles for old mediators
Venkatakrishna R. Jala and Bodduluri Haribabu, *Trends in Immunology* 25, 315–322

VEGF: necessary to prevent motoneuron degeneration, sufficient to treat ALS?
Diether Lambrechts, Erik Storkebaum and Peter Carmeliet, *Trends in Moelcular Medicine* 10, 275–282

Compound lipophilicity for substrate binding to human P450s in drug metabolism
David F.V. Lewis, Miriam N. Jacobs and Maurice Dickins, *Drug Discovery Today* 9, 530–537

Evaluating the validity of animal models for research into therapies for immune-based disorders
Bert A. 't Hart, Sandra Amor and Margreet Jonker, *Drug Discovery Today* 9, 517–524

Appendix IV

- Help
- General Help
- Formats
- Gaps
- Matrix
- References
- ClustalW Help
- ClustalW FAQ
- Jalview Help
- Scores Table
- Alignment
- Guide Tree
- Colours

ClustalW **Results**

Results of search	
Number of sequences	3
Alignment score	12762
Sequence format	Pearson
Sequence type	aa
ClustalW version	1.82
JalView	
Output file	clustalw-20050728-18060107.output
Alignment file	clustalw-20050728-18060107.aln
Guide tree file	clustalw-20050728-18060107.dnd
Your input file	clustalw-20050728-18060107.input
SUBMIT ANOTHER JOB	

To save a result file right-click the file link in the above table and choose "Save Target As".
If you cannot see the JalView button, reload the page and check your browser settings to enable Java Applets.

Scores Table

Sort by | Sequence Number | View Output File

SeqA	Name	Len(aa)	SeqB	Name	Len(aa)	Score
1	SEQ_ID_NO_7	842	2	SEQ_ID_NO_8	843	90
1	SEQ_ID_NO_7	842	3	SEQ_ID_NO_9	838	70
2	SEQ_ID_NO_8	843	3	SEQ_ID_NO_9	838	69

PLEASE NOTE: Some scores may be missing from the above table if the alignment was done using multiple CPU mode. Plea

Sort by | Sequence Number | View Output File

Alignment

Show Colors | View Alignment File

CLUSTAL W (1.82) multiple sequence alignment

```
SEQ_ID_NO_7       MGPQARTLCLLSLLLHVLPKPGKLVENSDFHLAGDYLLGGLFTLHANVKSISHLSYLQVP 60
SEQ_ID_NO_8       MGPQARTLHLLFLLLHALPKPVMLVGNSDFHLAGDYLLGGLFTLHANVKSVSHLSYLQVP 60
SEQ_ID_NO_9       MGPRAKTICSLFFLLWVLAEP---AENSDFYLPGDYLLGGLFSLHANMKGIVHLNFLQVP 57
                  ***:*:*:  * :** .*.:*   . ****:*.*********:****:*.: **.:****

SEQ_ID_NO_7       KCNEFTMKVLGYNLMQAMRFAVEEINNCSSLLPGVLLGYEMVDVCYLSNNIHPGLYFLAQ 120
SEQ_ID_NO_8       KCNEYNMKVLGYNLMQAMRFAVEEINNCSSLLPGVLLGYEMVDVCYLSNNIQPGLYFLSQ 120
SEQ_ID_NO_9       MCKEYEVKVIGYNLMQAMRFAVEEINNDSSLLPGVLLGYEIVDVCYISNNVQPVLYFLAH 117
                   *:*: :**:***************** ***********:*****:***::* ****::

SEQ_ID_NO_7       DDDLLPILKDYSQYMPHVVAVIGPDNSESAITVSNILSHFLIPQITYSAISDKLRDKRHF 180
SEQ_ID_NO_8       IDDFLPILKDYSQYRPQVVAVIGPDNSESAITVSNILSYFLVPQVTYSAITDKLQDKRRF 180
SEQ_ID_NO_9       EDNLLPIQEDYSNYISRVVAVIGPDNSESVMTVANFLSLFLLPQITYSAISDELRDKVRF 177
                   *::*** :***:* .:************.:**:*:** **:**:*****:*:*:** :*

SEQ_ID_NO_7       PSMLRTVPSATHHIEAMVQLMVHFQWNWIVVLVSDDDYGRENSHLLSQRLTKTSDICIAF 240
SEQ_ID_NO_8       PAMLRTVPSATHHIEAMVQLMVHFQWNWIVVLVSDDDYGRENSHLLSQRLTNTGDICIAF 240
SEQ_ID_NO_9       PALLRTTPSADHHVEAMVQLMLHFRWNWIIVLVSSDTYGRDNGQLLGERVARR-DICIAF 236
                  *::***.*** **:*******:**:****:****.* ***:*.:**.:*::.  ******

SEQ_ID_NO_7       QEVLPIPESSQVMRSEEQRQLDNILDKLRRTSARVVVVFSPELSLYSFFHEVLRWNFTGF 300
SEQ_ID_NO_8       QEVLPVPEPNQAVRPEEQDQLDNILDKLRRTSARVVVIFSPELSLHNFFREVLRWNFTGF 300
SEQ_ID_NO_9       QETLPTLQPNQNMTSEERQRLVTIVDKLQQSTARVVVVFSPDLTLYHFFNEVLRQNFTGA 296
                  **.**  :..* : .**: :* .*:***::::*****:***:*:*: **.**** ****

SEQ_ID_NO_7       VWIASESWAIDPVLHNLTELRHTGTFLGVTIQRVSIPGFSQFRVRRDKPGYPVPNTTNLR 360
SEQ_ID_NO_8       VWIASESWAIDPVLHNLTELRHTGTFLGVTIQRVSIPGFSQFRVRHDKPGYRMPNETSLR 360
SEQ_ID_NO_9       VWIASESWAIDPVLHNLTELGHLGTFLGITIQSVPIPGFSEFREWGPQAGPPPLSRTSQS 356
                  ******************** * *****:*** *.*****:**    :.*    . *.

SEQ_ID_NO_7       TTCNQDCDACLNTTKSFNNILILSGERVVYSVYSAVYAVAHALHRLLGCNRVRCTKQKVY 420
SEQ_ID_NO_8       TTCNQDCDACMNITESFNNVLMLSGERVVYSVYSAVYAVAHTLHRLLHCNQVRCTKQIVY 420
SEQ_ID_NO_9       YTCNQECDNCLNATLSFNTILRLSGERVVYSVYSAVYAVAHALHSLLGCDKSTCTKRVVY 416
                   ****:** *:* * ***.:* ********************:** ** *::  ***: **

SEQ_ID_NO_7       PWQLLREIWHVNFTLLGNRLFFDQQGDMPMLLDIIQWQWDLSQNPFQSIASYSPTSKRLT 480
SEQ_ID_NO_8       PWQLLREIWHVNFTLLGNQLFFDEQGDMPMLLDIIQWQWGLSQNPFQSIASYSPTETRLT 480
SEQ_ID_NO_9       PWQLLEEIWKVNFTLLDHQIFFDPQGDVALHLEIVQWQWDRSQNPFQSVASYYPLQRQLK 476
                  *****.***:******.:::*** ***:.: *:*:****. *******:*** * . :*.

SEQ_ID_NO_7       YINNVSWYTPNNTVPVSMCSKSCQPGQMKKSVGLHPCCFECLDCMPGTYLNRSADEFNCL 540
SEQ_ID_NO_8       YISNVSWYTPNNTVPISMCSKSCQPGQMKKPIGLHPCCFECVDCPPDTYLNRSVDEFNCL 540
SEQ_ID_NO_9       NIQDISWHTVNNTIPMSMCSKRCQSGQKKKPVGIHVCCFECIDCLPGTFLNHTEDEYECQ 536
                   *.::**:* ***:*:***** **.** **.:*:* *****:** *.*:**:: **::*

SEQ_ID_NO_7       SCPGSMWSYKNDITCFQRRPTFLEWHEVPTIVVAILAALGFFSTLAILFIFWRHFQTPMV 600
SEQ_ID_NO_8       SCPGSMWSYKNNIACFKRRLAFLEWHEVPTIVVTILAALGFISTLAILLIFWRHFQTPMV 600
SEQ_ID_NO_9       ACPNNEWSYQSETSCFKRQLVFLEWHEAPTIAVALLAALGFLSTLAILVIFWRHFQTPIV 596
                  :**.. ***:.: :**:*: .******.***.*::******:******.*********:*

SEQ_ID_NO_7       RSAGGPMCFLMLVPLLLAFGMVPVYVGPPTVFSCFCRQAFFTVCFSICLSCITVRSFQIV 660
SEQ_ID_NO_8       RSAGGPMCFLMLVPLLLAFGMVPVYVGPPTVFSCFCRQAFFTVCFSVCLSCITVRSFQIV 660
SEQ_ID_NO_9       RSAGGPMCFLMLTLLLVAYMVVPVYVGPPKVSTCLCRQALFPLCFTICISCIAVRSFQIV 656
                  ************. **:*:.:********.* :*:****:*.:**::*:***:*******

SEQ_ID_NO_7       CVFKMARRLPSAYSFWMRYHGPYVFVAFITAIKVALVVGNMLATTINPIGRTDPDDPNIM 720
SEQ_ID_NO_8       CVFKMARRLPSAYGFWMRYHGPYVFVAFITAVKVALVAGNMLATTINPIGRTDPDDPNII 720
SEQ_ID_NO_9       CAFKMASRFPRAYSYWVRYQGPYVSMAFITVLKMVIVVIGMLARPQSHP-RTDPDDPKIT 715
                  *.**** *:* **.:*:**:**** :****.:*:.:*. .*** . .   ********:*

SEQ_ID_NO_7       ILSCHPNYRNGLLFNTSMDLLLSVLGFSFAYMGKELPTNYNEAKFITLSMTFSFTSSISL 780
SEQ_ID_NO_8       ILSCHPNYRNGLLFNTSMDLLLSVLGFSFAYVGKELPTNYNEAKFITLSMTFSFTSSISL 780
SEQ_ID_NO_9       IVSCNPNYRNSLLFNTSLDLLLSVVGFSFAYMGKELPTNYNEAKFITLSMTFYFTSSVSL 775
                  *:**:*****.******:******:******:********************* ****:**

SEQ_ID_NO_7       CTFMSVHDGVLVTIMDLLVTVLNFLAIGLGYFGPKCYMILFYPERNTSAYFNSMIQGYTM 840
SEQ_ID_NO_8       CTFMSVHDGVLVTIMDLLVTVLNFLAIGLGYFGPKCYMILFYPERNTSAYFNSMIQGYTM 840
SEQ_ID_NO_9       CTFMSAYSGVLVTIVDLLVTVLNLLAISLGYFGPKCYMILFYPERNTPAYFNSMIQGYTM 835
                  *****.:.******:********:***.******************.************

SEQ_ID_NO_7       RK- 842
```

```
SEQ_ID_NO_8       RKS 843
SEQ_ID_NO_9       RRD 838
                  *:
```

PLEASE NOTE: Showing colors on large alignments is slow.

Show Colors | View Alignment File

Guide Tree

Show as Phylogram Tree | Show Distances | View DND File

```
(
SEQ_ID_NO_7:0.03678,
SEQ_ID_NO_8:0.05348,
SEQ_ID_NO_9:0.25439);
```

Cladogram

Show as Phylogram Tree | Show Distances | View DND File

Right-click on the above tree to see display options.

Problems printing? Read how to print a Phylogram or Cladogram.

Please contact EBI Support with any problems or suggestions regarding this site.

View Printer-friendly version of this page | Terms of Use

Appendix V

- Help
- General Help
- Formats
- Gaps
- Matrix
- References
- ClustalW Help
- ClustalW FAQ
- Jalview Help
- Scores Table
- Alignment
- Guide Tree
- Colours

ClustalW **Results**

Results of search	
Number of sequences	4
Alignment score	27703
Sequence format	Pearson
Sequence type	aa
ClustalW version	1.82
JalView	
Output file	clustalw-20050728-18221731.output
Alignment file	clustalw-20050728-18221731.aln
Guide tree file	clustalw-20050728-18221731.dnd
Your input file	clustalw-20050728-18221731.input
SUBMIT ANOTHER JOB	

To save a result file right-click the file link in the above table and choose "Save Target As".
If you cannot see the JalView button, reload the page and check your browser settings to enable Java Applets.

Scores Table

Sort by | Sequence Number | View Output File

SeqA	Name	Len(aa)	SeqB	Name	Len(aa)	Score
1	SEQ_ID_NO_15	852	2	SEQ_ID_NO_20	858	72
1	SEQ_ID_NO_15	852	3	SEQ_ID_NO_23	857	72
1	SEQ_ID_NO_15	852	4	SEQ_ID_NO_25	858	73
2	SEQ_ID_NO_20	858	3	SEQ_ID_NO_23	857	99
2	SEQ_ID_NO_20	858	4	SEQ_ID_NO_25	858	92
3	SEQ_ID_NO_23	857	4	SEQ_ID_NO_25	858	92

PLEASE NOTE: Some scores may be missing from the above table if the alignment was done using multiple CPU mode. Plea

Sort by | Sequence Number | View Output File

Alignment

Show Colors | View Alignment File

CLUSTAL W (1.82) multiple sequence alignment

```
SEQ_ID_NO_15    MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
SEQ_ID_NO_20    MPALAIMGLSLAAFLELGMGASLCLSQQFKAQGDYILGGLFPLGSTEEATLNQRTQPNSI 60
SEQ_ID_NO_23    MPALAIMGLSLAAFLELGMGASLCLSQQFKAQGDYILGGLFPLGSTEEATLNQRTQPNSI 60
SEQ_ID_NO_25    MPGLAILGLSLAAFLELGMGSSLCLSQQFKAQGDYILGGLFPLGTTEEATLNQRTQPNGI 60
                * . *::**** *:*. * *:.******:: :***:******** :*** *..**:*..

SEQ_ID_NO_15    VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
SEQ_ID_NO_20    LCNRFSPLGLFLAMAMKMAVEEINNGSALLPGLRLGYDLFDTCSEPVVTMKSSLMFLAKV 120
SEQ_ID_NO_23    PCNRFSPLGLFLAMAMKMAVEEINNGSALLPGLRLGYDLFDTCSEPVVTMKSSLMFLAKV 120
SEQ_ID_NO_25    LCTRFSPLGLFLAMAMKMAVEEINNGSALLPGLRLGYDLFDTCSEPVVTMKPSLMFMAKV 120
                 *.***. **: *:*********** * ********************:**.****:**.

SEQ_ID_NO_15    GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
SEQ_ID_NO_20    GSQSIAAYCNYTQYQPRVLAVIGPHSSELALITGKFFSFFLMPQVSYSASMDRLSDRETF 180
SEQ_ID_NO_23    GSQSIAAYCNYTQYQPRVLAVIGPHSSELALITGKFFSFFLMPQVSYSASMDRLSDRETF 180
SEQ_ID_NO_25    GSQSIAAYCNYTQYQPRVLAVIGPHSSELALITGKFFSFFLMPQVSYSASMDRLSDRETF 180
                **:.**************************::***************.***: ** ****

SEQ_ID_NO_15    PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
SEQ_ID_NO_20    PSFFRTVPSDRVQLQAVVTLLQNFSWNWVAALGSDDDYGREGLSIFSSLANARGICIAHE 240
SEQ_ID_NO_23    PSFFRTVPSDRVQLQAVVTLLQNFSWNWVAALGSDDDYGREGLSIFSSLANARGICIAHE 240
SEQ_ID_NO_25    PSFFRTVPSDRVQLQAVVTLLQNFSWNWVAALGSDDDYGREGLSIFSGLANSRGICIAHE 240
                ************** *.. ***:*.***********:***:******.** :********

SEQ_ID_NO_15    GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
SEQ_ID_NO_20    GLVPQHDTSGQQLGKVLDVLCQVNQSKVQVVVLFASARAVYSLFSYSIHHGLSPKVWVAS 300
SEQ_ID_NO_23    GLVPQHDTSGQQLGKVLDVLRQVNQSKVQVVVLFASARAVYSLFSYSIHHGLSPKVWVAS 300
SEQ_ID_NO_25    GLVPQHDTSGQQLGKVVDVLRQVNQSKVQVVVLFASARAVYSLFSYSILHDLSPKVWVAS 300
                ****   :...:**** *** *****.****:****.:*.::**.***   *********

SEQ_ID_NO_15    EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
SEQ_ID_NO_20    ESWLTSDLVMTLPNIARVGTVLGFLQRGALLPEFSHYVETHLALAADPAFCASLN-AELD 359
SEQ_ID_NO_23    ESWLTSDLVMTLPNIARVGTVLGFLQRGALLPEFSHYVETHLALAADPAFCASLN-AELD 359
SEQ_ID_NO_25    ESWLTSDLVMTLPNIARVGTVLGFLQRGALLPEFSHYVETRLALAADPTFCASLK-AELD 359
                *:******** **.:*::*********** * **.:**:*:****:**:**::*   * .

SEQ_ID_NO_15    LEEDVVGQRCPQCDCITLQNVSAGLN--------HHQTFSVYAAVYSVAQALHNTLQCNA 412
SEQ_ID_NO_20    LEEHVMGQRCPQCDDIMLQNLSSGLLQNLSAGQLHHQIFATYAAVYSVAQALHNTLQCNV 419
SEQ_ID_NO_23    LEEHVMGQRCPRCDDIMLQNLSSGLLQNLSAGQLHHQIFATYAAVYSVAQALHNTLQCNV 419
SEQ_ID_NO_25    LEERVMGPRCSQCDYIMLQNLSSGLMQNLSAGQLHHQIFATYAAVYSVAQALHNTLQCNV 419
                *** *:* **.:** * ***:*:**         *** *:.******************.

SEQ_ID_NO_15    SGCPAQDPVKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVG 472
SEQ_ID_NO_20    SHCHVSEHVLPWQLLENMYNMSFHARDLTLQFDAEGNVDMEYDLKMWVWQSPTPVLHTVG 479
SEQ_ID_NO_23    SHCHVSEHVLPWQLLENMYNMSFHARDLTLQFDAEGNVDMEYDLKMWVWQSPTPVLHTVG 479
SEQ_ID_NO_25    SHCHTSEPVQPWQLLENMYNMSFRARDLTLQFDAKGSVDMEYDLKMWVWQSPTPVLHTVG 479
                * * ..: * **********::*:. .*.*:**:.*.********:****...* ** **

SEQ_ID_NO_15    RFNGSLRTERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQN 532
SEQ_ID_NO_20    TFNGTLQLQQSKMYWPG--NQVPVSQCSRQCKDGQVRRVKGFHSCCYDCVDCKAGSYRKH 537
SEQ_ID_NO_23    TFNGTLQLQQSKMYWPG--NQVPVSQCSRQCKDGQVRRVKGFHSCCYDCVDCKAGSYRKH 537
SEQ_ID_NO_25    TFNGTLQLQHSKMYWPG--NQVPVSQCSRQCKDGQVRRVKGFHSCCYDCVDCKAGSYRKH 537
                 ***:*: :: *: *    ** ***:*****::******************:*****::

SEQ_ID_NO_15    PDDIACTFCGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVH 592
SEQ_ID_NO_20    PDDFTCTPCNQDQWSPEKSTACLPRRPKFLAWGEPVVLSLLLLLCLVLGLALAALGLSVH 597
SEQ_ID_NO_23    PDDFTCTPCNQDQWSPEKSTACLPRRPKFLAWGEPVVLSLLLLLCLVLGLALAALGLSVH 597
SEQ_ID_NO_25    PDDFTCTPCGKDQWSPEKSTTCLPRRPKFLAWGEPAVLSLLLLLCLVLGLTLAALGLFVH 597
                ***::** *.:*:****:** *: **.:********.** *****.*.***.****** **

SEQ_ID_NO_15    HRDSPLVQASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLF 652
SEQ_ID_NO_20    HWDSPLVQASGGSQFCFGLICLGLFCLSVLLFPGRPSSASCLAQQPMAHLPLTGCLSTLF 657
SEQ_ID_NO_23    HWDSPLVQASGGSQFCFGLICLGLFCLSVLLFPGRPSSASCLAQQPMAHLPLTGCLSTLF 657
SEQ_ID_NO_25    YWDSPLVQASGGSLFCFGLICLGLFCLSVLLFPGRPRSASCLAQQPMAHLPLTGCLSTLF 657
                : **********.  ****:****.*********:* .* ******::************
```

```
SEQ_ID_NO_15      LQAAEIFVESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWH 712
SEQ_ID_NO_20      LQAAETFVESELPLSWANWLCSYLRGLWAWLVVLSATFVEAALCAWYLTAFPPEVVTDWS 717
SEQ_ID_NO_23      LQAAETFVESELPLSWANWLCSYLRGLWAWLVVLLATFVEAALCAWYLIAFPPEVVTDWS 717
SEQ_ID_NO_25      LQAAEIFVESELPLSWANWLCSYLRGPWAWLVVLLATLVEAALCAWYLMAFPPEVVTDWQ 717
                  ***** ***********: *.. *** ******* * :**.***:*** **********

SEQ_ID_NO_15      MLPTEALVHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYF 772
SEQ_ID_NO_20      VLPTEVLEHCHVRSWVSLGLVHITNAMLAFLCFLGTFLVQSQPGRYNRARGLTFAMLAYF 777
SEQ_ID_NO_23      VLPTEVLEHCHVRSWVSLGLVHITNAMLAFLCFLGTFLVQSQPGRYNRARGLTFAMLAYF 777
SEQ_ID_NO_25      VLPTEVLEHCRMRSWVSLGLVHITNAVLAFLCFLGTFLVQSQPGRYNRARGLTFAMLAYF 777
                  :****.* **: .*****:**.* *** ************:**** ***************

SEQ_ID_NO_15      ITWVSFVPLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGG 832
SEQ_ID_NO_20      ITWVSFVPLLANVQVAYQPAVQMGAILVCALGILVTFHLPKCYVLLWLPKLNTQEFFLGR 837
SEQ_ID_NO_23      ITWVSFVPLLANVQVAYQPAVQMGAILVCALGILVTFHLPKCYVLLWLPKLNTQEFFLGR 837
SEQ_ID_NO_25      IIWVSFVPLLANVQVAYQPAVQMGAILFCALGILATFHLPKCYVLLWLPELNTQEFFLGR 837
                  * ************. :*******:*.*.****.:****:**:*:  * *** *****

SEQ_ID_NO_15      GPGDAQGQNDGNTGNQGKHE- 852
SEQ_ID_NO_20      NAKKAADENSGGGEAAQEHNE 858
SEQ_ID_NO_23      NAKKAADENSGGGEAAQGHN- 857
SEQ_ID_NO_25      SPKEASDGNSGSSEATRGHSE 858
                  .. .* . *.*.      *.
```

PLEASE NOTE: Showing colors on large alignments is slow.

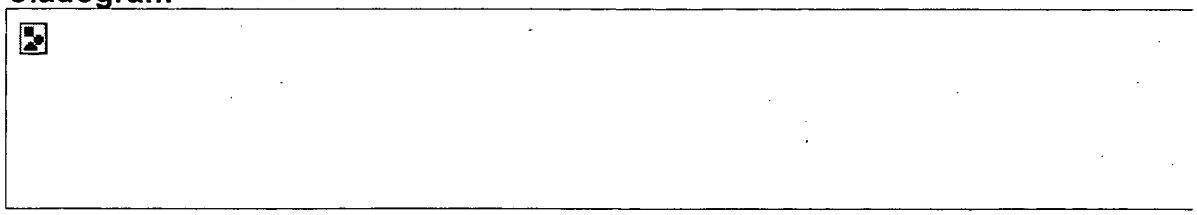
Show Colors | View Alignment File

Guide Tree

Show as Phylogram Tree | Show Distances | View DND File

```
(
SEQ_ID_NO_15:0.23527,
(
SEQ_ID_NO_20:0.00377,
SEQ_ID_NO_23:0.00323)
:0.03353,
SEQ_ID_NO_25:0.03469);
```

Cladogram



Show as Phylogram Tree | Show Distances | View DND File

Right-click on the above tree to see display options.

Problems printing? Read how to print a Phylogram or Cladogram.

Please contact EBI Support with any problems or suggestions regarding this site.

View Printer-friendly version of this page | Terms of Use

Appendix VI

Help **ClustalW**

LAUNCH APPLICATION

• INTRODUCTION

Multiple alignments of protein sequences are important tools in studying sequences. The basic information they provide is identification of conserved sequence regions. This is very useful in designing experiments to test and modify the function of specific proteins, in predicting the function and structure of proteins, and in identifying new members of protein families.

Sequences can be aligned across their entire length (global alignment) or only in certain regions (local alignment). This is true for pairwise and multiple alignments. Global alignments need to use gaps (representing insertions/deletions) while local alignments can avoid them, aligning regions between gaps. ClustalW is a fully automatic program for global multiple alignment of DNA and protein sequences. The alignment is progressive and considers the sequence redundancy. Trees can also be calculated from multiple alignments. The program has some adjustable parameters with reasonable defaults.

ClustalW Nucleotide Tutorial
ClustalW Protein Tutorial

For additional help on ClustalW also see:
- **ClustalW FAQ**
- **ClustalW Improving Sensitivity**

• YOUR SEQUENCES

Please make sure that your sequences have different names as the first 30 characters of the name are significant. If clustalw finds two or more sequences with the same name it will fail! View example. Click here for more information on sequence errors.

ClustalW currently supports 7 multiple sequence formats. These are:

- NBRF/PIR
- EMBL / UniProtKB/Swiss-Prot
- Pearson (Fasta)
- GDE
- ALN/ClustalW
- GCG/MSF
- RSF

More on Sequence Formats

Please remove any white space space or empty lines from the beginning of your input.

N.B. Bootstrapping is possible, you can paste your .aln files from your results back into the clustalw submission form.

• YOUR EMAIL

You must type your email address in this text box if you are running a job via email. It is not necessary to fill in the box if you are running your search interactively.

• ALIGNMENT TITLE

You may type any text you want to help you identify your search results.

- **RESULTS**

This option lets you choose between email and interactive runs. The email run requires you to type an email address in the email text box, and your results will be delivered when they are ready to your email address, thus avoiding waiting for your results as with an interactive run. For example: joe@somewhere.domain.country.

The default is interactive.

- **ALIGNMENT**

You may choose to run a full alignment or using a stringent algorithm for generating the tree guide or a fast algorithm.

- **CPU MODE**

The multiple CPU option run a special version of clustalw using several linux pc nodes in a parallel fashion to increase the speed of the job without compromising the quality of the results. This option is to be chosen when the user has a large number of sequences (50+ but less than 500) to align. However, care should be taken not to overestimate the quantification of the results. A very large alignment is difficult to read and handle by other software.

- **OUTPUT**

Here you decide which output format you want your multiple sequence alignment in. The options are ALN, GCG, PHYLIP, PIR and GDE. The output results are stored in a ".aln" file which contains your results in ALN format.

You can configure your browser to automatically load the results files from clustalw into an suitable application of your choice. Here is a small list of URL's for obtaining such applications for Win95, Macs and UNIX boxes:

- **belvu - UNIX Multiple sequence alignment viewer -**
  http://www.cgr.ki.se/cgr/groups/sonnhammer/Belvu.html
- **njplot - UNIX, Mac & PC - Tree viewer -**
  http://pbil.univ-lyon1.fr/software/njplot.html
- GeneDoc - GCG MSF file viewer -
  (NB - Use GCG format in the OUTPUT option) -
  http://www.psc.edu/biomed/genedoc/
- **TreeView - Tree viewer for Macs and PC (running Windows)**
  http://taxonomy.zoology.gla.ac.uk/rod/treeview.html
- **ClustalX - Graphical Interface X-Windows/Mac/Win9x based version of ClustalW**
  ftp://ftp.ebi.ac.uk/pub/software/dos/clustalx/
  ftp://ftp.ebi.ac.uk/pub/software/mac/clustalx/
  ftp://ftp.ebi.ac.uk/pub/software/unix/clustalx/
- **ClustalW - Command line version**
  ftp://ftp.ebi.ac.uk/pub/software/dos/clustalw/
  ftp://ftp.ebi.ac.uk/pub/software/mac/clustalw/
  ftp://ftp.ebi.ac.uk/pub/software/unix/clustalw/
- **CINEMA -**
  A Colour INteractive Editor for Multiple Alignments
  http://www.bioinf.man.ac.uk/dbbrowser/CINEMA2.1/
- **EMMA-**
  An EMBOSS interface to clustalw.
  http://www.emboss.org/

To configure your browser you must tell it how to handle the following non-registered MIME types:

- application/x-tree njplot %s (or whatever application you choose to handle the MIME type)

- application/x-align belvu %s (or whatever application you choose to handle the MIME type)

JalView: A new experimental option has been added to the results page which involved using a Java Applet called JalView. This is a fully featured MSA editor which allows you not only to edit the alignment but also, to exchange the alignment formats. Please note that JalView is under development. For documentation please click on the JalView Hyperlink or visit the official website at http://www.jalview.org.

- **OUTORDER**

Decide which order the sequences should be printed in the alignment.

- **COLOUR**

Show Colors

A button labeled 'Show Colors' will be displayed in the Alignment section of results page. If you press this button the alignment will be show in color according to the table below.

NOTE: NOTE: This option only works when you have chosen ALN or GCG as the output format.

AVFPMILW	RED	Small (small+ hydrophobic (incl.aromatic -Y))
DE	BLUE	Acidic
RHK	MAGENTA	Basic
STYHCNGQ	GREEN	Hydroxyl + Amine + Basic - Q
Others	Gray	

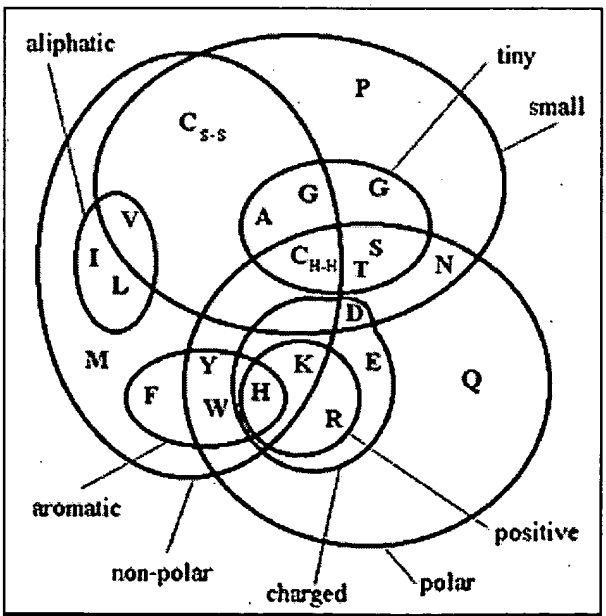


Figure 1. A Venn diagram showing the relationship of the 20 naturally occuring amino acids to a selection of physio-chemical properties thought to be important in the determination of protein structure.

Hide Colors

A button labeled 'Hide Colors' will be displayed in the result page is the 'Show

colors' button has been pressed. This option will display the alignment in normal text with no colors.

The default is not to colour the alignments.

CONSENSUS SYMBOLS:

An alignment will display by default the following symbols denoting the degree of conservation observed in each column:

"*" means that the residues or nucleotides in that column are identical in all sequences in the alignment.

":" means that conserved substitutions have been observed, according to the COLOUR table above.

"." means that semi-conserved substitutions are observed.

More on sequence/alignment formats

FAST PAIRWISE ALIGNMENT OPTIONS:

- **KTUP**
  This option allows you to choose which 'word-length' to use when calculating fast pairwise alignments.(note: make sure you have chosen 'fast' in the ALIGNMENT.

- **WINDOW**
  Use this option to set the window length when calculating fast pairwise alignments.(Note: make sure you have chosen 'fast' in the ALIGNMENT.

- **SCORE**
  This option allows you to decide which score to take into account when calculating a fast pairwise alignment. (Note: make sure you have chosen 'fast' in the ALIGNMENT.

- **TOPDIAG**
  Select here how many top diagonals should be integrated when calculating a fast pairwise alignment.(Note: make sure you have chosen 'fast' in the ALIGNMENT.

- **PAIRGAP**
  Select here to set the gap penalty when generating fast pairwise alignments.

MULTIPLE SEQUENCE ALIGNMENT OPTIONS:

- **MATRIX**
  This option allows you to choose which matrix series to use when generating the multiple sequence alignment. The program goes through the chosen matrix series, spanning the full range of amino acid distances.
  - *BLOSUM (Henikoff).* These matrices appear to be the best available for carrying out data base similarity searches.
    The matrix used is Blosum30.
  - *PAM (Dayhoff).* These have been extremely widely used since the late '70s. We use the PAM 350 matrix.
  - *GONNET.* These matrices were derived using almost the same procedure as the Dayhoff one (above) but are much more up to date

and are based on a far larger data set. They appear to be more sensitive than the Dayhoff series. We use the GONNET 250 matrix.

We also supply an identity matrix which gives a score of 10 to two identical amino acids and a score of zero otherwise.

- Default values are:
  - ***DNA:*** *DNA Identity matrix.*
  - ***Protein:*** *Gonnet 250.*

More on Matrices

- **GAPOPEN**
  You can set here the penalty for opening a gap.

  Default values are:
  - ***DNA:*** *15.0*
  - ***Protein:*** 10.0

- **ENDGAP**
  You can set here the penalty for closing a gap.

  Default value is -1.

- **GAPEXT**
  You can set here the penalty for extending a gap.

  Default values are:
  - ***DNA:*** *6.66*
  - ***Protein:*** 0.2

- **GAPDIST**
  You can set here the gap separation penalty.

  Default value is 4.

More about gaps.

- **EXAMPLE**

e.g. A multiple sequence alignment was done with ClustalW. 5 sequences were input in the fasta format (Download):

```
>FOSB_MOUSE Protein fosB
     MFQAFPGDYD SGSRCSSSPS AESQYLSSVD SFGSPPTAAA SQECAGLGEM PGSFVPTVTA
     ITTSQDLQWL VQPTLISSMA QSQGQPLASQ PPAVDPYDMP GTSYSTPGLS AYSTGGASGS
     GGPSTSTTTS GPVSARPARA RPRRPREETL TPEEEEKRRV RRERNKLAAA KCRNRRRELT
     DRLQAETDQL EEEKAELESE IAELQKEKER LEFVLVAHKP GCKIPYEEGP GPGPLAEVRD
     LPGSTSAKED GFGWLLPPPP PPPLPFQSSR DAPPNLTASL FTHSEVQVLG DPFPVVSPSY
     TSSFVLTCPE VSAFAGAQRT SGSEQPSDPL NSPSLLAL

>FOSB_HUMAN Protein fosB
     MFQAFPGDYD SGSRCSSSPS AESQYLSSVD SFGSPPTAAA SQECAGLGEM PGSFVPTVTA
     ITTSQDLQWL VQPTLISSMA QSQGQPLASQ PPVVDPYDMP GTSYSTPGMS GYSSGGASGS
     GGPSTSGTTS GPGPARPARA RPRRPREETL TPEEEEKRRV RRERNKLAAA KCRNRRRELT
     DRLQAETDQL EEEKAELESE IAELQKEKER LEFVLVAHKP GCKIPYEEGP GPGPLAEVRD
     LPGSAPAKED GFSWLLPPPP PPPLPFQTSQ DAPPNLTASL FTHSEVQVLG DPFPVVNPSY
     TSSFVLTCPE VSAFAGAQRT SGSDQPSDPL NSPSLLAL

>FOS_CHICK Proto-oncogene protein c-fos
     MMYQGFAGEY EAPSSRCSSA SPAGDSLTYY PSPADSFSSM GSPVNSQDFC TDLAVSSANF
     VPTVTAISTS PDLQWLVQPT LISSVAPSQN RGHPYGVPAP APPAAYSRPA VLKAPGGRGQ
     SIGRRGKVEQ LSPEEEEKRR IRRERNKMAA AKCRNRRREL TDTLQAETDQ LEEEKSALQA
     EIANLLKEKE KLEFILAAHR PACKMPEELR FSEELAAATA LDLGAPSPAA AEEAFALPLM
     TEAPPAVPPK EPSGSGLELK AEPFDELLFS AGPREASRSV PDMDLPGASS FYASDWEPLG
     AGSGGELEPL CTPVVTCTPC PSTYTSTFVF TYPEADAFPS CAAAHRKGSS SNEPSSDSLS
```

```
     SPTLLAL

>FOS_RAT Proto-oncogene protein c-fos
     MMFSGFNADY EASSSRCSSA SPAGDSLSYY HSPADSFSSM GSPVNTQDFC ADLSVSSANF
     IPTVTAISTS PDLQWLVQPT LVSSVAPSQT RAPHPYGLPT PSTGAYARAG VVKTMSGGRA
     QSIGRRGKVE QLSPEEEEKR RIRRERNKMA AAKCRNRRRE LTDTLQAETD QLEDEKSALQ
     TEIANLLKEK EKLEFILAAH RPACKIPNDL GFPEEMSVTS LDLTGGLPEA TTPESEEAFT
     LPLLNDPEPK PSLEPVKNIS NMELKAEPFD DFLFPASSRP SGSETARSVP DVDLSGSFYA
     ADWEPLHSSS LGMGPMVTEL EPLCTPVVTC TPSCTTYTSS FVFTYPEADS FPSCAAAHRK
     GSSSNEPSSD SLSSPTLLAL

>FOS_MOUSE Proto-oncogene protein c-fos
     MMFSGFNADY EASSSRCSSA SPAGDSLSYY HSPADSFSSM GSPVNTQDFC ADLSVSSANF
     IPTVTAISTS PDLQWLVQPT LVSSVAPSQT RAPHPYGLPT QSAGAYARAG MVKTVSGGRA
     QSIGRRGKVE QLSPEEEEKR RIRRERNKMA AAKCRNRRRE LTDTLQAETD QLEDEKSALQ
     TEIANLLKEK EKLEFILAAH RPACKIPDDL GFPEEMSVAS LDLTGGLPEA STPESEEAFT
     LPLLNDPEPK PSLEPVKSIS NVELKAEPFD DFLFPASSRP SGSETSRSVP DVDLSGSFYA
     ADWEPLHSNS LGMGPMVTEL EPLCTPVVTC TPGCTTYTSS FVFTYPEADS FPSCAAAHRK
     GSSSNEPSSD SLSSPTLLAL
```

Output was in the format:

The sequences are aligned with each other, with the query sequence at the top and subsequent sequences below. Gaps are represented by the "-" symbol. The running total number of amino acids or nucleotides are shown on the right. The degree of similarity is illustrated underneath the alignments with a series of consensus symbols.

Consensus Symbols

An alignment will display by default the following symbols denoting the degree of conservation observed in each column:

"*" means that the residues or nucleotides in that column are identical in all sequences in the alignment.

":" means that conserved substitutions have been observed, according to the COLOUR table above.

"." means that semi-conserved substitutions are observed.

```
FOS_RAT         MMFSGFNADYEASSSRCSSASPAGDSLSYYHSPADSFSSMGSPVNTQDFCADLSVSSANF 60
FOS_MOUSE       MMFSGFNADYEASSSRCSSASPAGDSLSYYHSPADSFSSMGSPVNTQDFCADLSVSSANF 60
FOS_CHICK       MMYQGFAGEYEAPSSRCSSASPAGDSLTYYPSPADSFSSMGSPVNSQDFCTDLAVSSANF 60
FOSB_MOUSE      -MFQAFPGDYDS-GSRCSS-SPSAESQ--YLSSVDSFGSPPTAAASQE-CAGLGEMPGSF 54
FOSB_HUMAN      -MFQAFPGDYDS-GSRCSS-SPSAESQ--YLSSVDSFGSPPTAAASQE-CAGLGEMPGSF 54
                 *:..* .:*:: .***** **:.:*   * *..***.*  :.. :*: *:.*.  ...*

FOS_RAT         IPTVTAISTSPDLQWLVQPTLVSSVAPSQ-------TRAPHPYGLPTPS-TGAYARAGVV 112
FOS_MOUSE       IPTVTAISTSPDLQWLVQPTLVSSVAPSQ-------TRAPHPYGLPTQS-AGAYARAGMV 112
FOS_CHICK       VPTVTAISTSPDLQWLVQPTLISSVAPSQ-------NRG-HPYGVPAPAPPAAYSRPAVL 112
FOSB_MOUSE      VPTVTAITTSQDLQWLVQPTLISSMAQSQGQPLASQPPAVDPYDMPGTS----YSTPGLS 110
FOSB_HUMAN      VPTVTAITTSQDLQWLVQPTLISSMAQSQGQPLASQPPVVDPYDMPGTS----YSTPGMS 110
                :******:** **********:**:* **... ::.    .**.:*  :    *: ..:

FOS_RAT         KTMSGGRAQSIG--------------------RRGKVEQLSPEEEEKRRIRRERNKMAAA 152
FOS_MOUSE       KTVSGGRAQSIG--------------------RRGKVEQLSPEEEEKRRIRRERNKMAAA 152
FOS_CHICK       KAP-GGRGQSIG--------------------RRGKVEQLSPEEEEKRRIRRERNKMAAA 151
FOSB_MOUSE      AYSTGGASGSGGPSTSTTTSGPVSARPARARPRRPREETLTPEEEEKRRVRRERNKLAAA 170
FOSB_HUMAN      GYSSGGASGSGGPSTSGTTSGPGPARPARARPRRPREETLTPEEEEKRRVRRERNKLAAA 170
                   :** . * *.::: :::.. .: .: : .** : * *:********:******:***

FOS_RAT         KCRNRRRELTDTLQAETDQLEDEKSALQTEIANLLKEKEKLEFILAAHRPACKIPNDLGF 212
FOS_MOUSE       KCRNRRRELTDTLQAETDQLEDEKSALQTEIANLLKEKEKLEFILAAHRPACKIPDDLGF 212
FOS_CHICK       KCRNRRRELTDTLQAETDQLEEEKSALQAEIANLLKEKEKLEFILAAHRPACKMPEELRF 211
FOSB_MOUSE      KCRNRRRELTDRLQAETDQLEEEKAELESEIAELQKEKERLEFVLVAHKPGCKIPYEEG- 229
FOSB_HUMAN      KCRNRRRELTDRLQAETDQLEEEKAELESEIAELQKEKERLEFVLVAHKPGCKIPYEEG- 229
                ********** *********:**: *::***:* ****:***:*.**:*.**:* :

FOS_RAT         PEEMSVTS-LDLTGGLPEATTPESEEAFTLPLLNDPEPK-PSLEPVKNISNMELKAEPFD 270
FOS_MOUSE       PEEMSVAS-LDLTGGLPEASTPESEEAFTLPLLNDPEPK-PSLEPVKSISNVELKAEPFD 270
FOS_CHICK       SEELAAATALDLG----APSPAAAEEAFALPLMTEAPPAVPPKEPSG--SGLELKAEPFD 265
```

```
FOSB_MOUSE      PGPGPLAEVRDLPG-----STSAKEDGFGWLLPPPPPPP-----------------LPFQ 267
FOSB_HUMAN      PGPGPLAEVRDLPG-----SAPAKEDGFSWLLPPPPPPP-----------------LPFQ 267
                .   . :   ** .     :..  *:.*   *   . *                   **:

FOS_RAT         DFLFPASSRPSGSETARSVPDVDLSG--SFYAADWEPLHSSSLGMGPMVTELEPLCTPVV 328
FOS_MOUSE       DFLFPASSRPSGSETSRSVPDVDLSG--SFYAADWEPLHSNSLGMGPMVTELEPLCTPVV 328
FOS_CHICK       ELLFSAGPR----EASRSVPDMDLPGASSFYASDWEPLGAGSGG------ELEPLCTPVV 315
FOSB_MOUSE      --------------SSRDAP-PNLTA--SLFTHS----------------EVQVLGDPFP 294
FOSB_HUMAN      --------------TSQDAP-PNLTA--SLFTHS----------------EVQVLGDPFP 294
                             :::..*  :*..  *::: .               *:: *  *.

FOS_RAT         TCTPSCTTYTSSFVFTYPEADSFPSCAAAHRKGSSSNEPSSDSLSSPTLLAL 380
FOS_MOUSE       TCTPGCTTYTSSFVFTYPEADSFPSCAAAHRKGSSSNEPSSDSLSSPTLLAL 380
FOS_CHICK       TCTPCPSTYTSTFVFTYPEADAFPSCAAAHRKGSSSNEPSSDSLSSPTLLAL 367
FOSB_MOUSE      VVSP---SYTSSFVLTCPEVSAF---AGAQR--TSGSEQPSDPLNSPSLLAL 338
FOSB_HUMAN      VVNP---SYTSSFVLTCPEVSAF---AGAQR--TSGSDQPSDPLNSPSLLAL 338
                . .*   :***:**:* **..:*   *.*:*  :*..: .**.*.**:****
```

• EDITING AN ALIGNMENT

You can edit the alignment using jalview. Click on the button below to view the above alignment.

• PHYLOGENETIC TREE

Phylogram is a branching diagram (tree) assumed to be an estimate of a phylogeny, branch lengths are proportional to the amount of inferred evolutionary change. A Cladogram is a branching diagram (tree) assumed to be an estimate of a phylogeny where the branches are of equal length, thus cladograms show common ancestry, but do not indicate the amount of evolutionary "time" separating taxa. Tree distances can be shown, just click on the diagram to get a menu of options. The ".dnd" file is a file that describes the phylogenetic tree.

These are now in controlled with new buttons in the output file as well as a pop up menu, that is available by right-clicking on the applet. The buttons on the page include "Show as Phylogram Tree", "Show as Cladogram Tree" and "Show Distances".

IMPORTANT!

Please note applets are not printed out with html pages, You will need to:

- Use the "Print Screen" button in the top right corner of your keyboard.
- Open an imaging application like paint or photoshop.
- Go "file>new" from the menu or "control+N" from the keyboard to create a new image.
- Go "edit>paste from the menu or "control+V" from the keyboard to paste your screen capture.
- The use the crop function to trim the image (e.g. "image>crop").
- Then save or print the image.

If you cannot see the tree, your java plug in is probably not working, please reinstall java on your machine or enable it in your internet settings. You can download java from http://java.sun.com/. If you have java running, you can see the version of java you are running in the box below (applet from http://www.javatester.org). You should have at least java version 1.2 to run this applet.

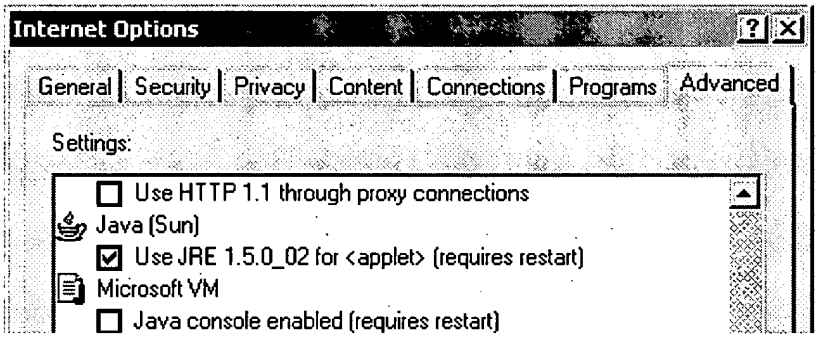
The vendor and version from the JVM:

You should see a pink box above with one line of text that says something like:

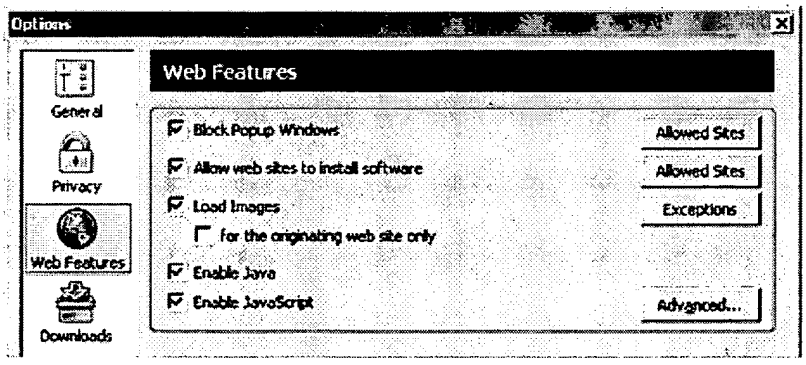
```
Java Version: 1.1.4 from Microsoft Corp.  or
Java Version 1.4.2_05 from Sun Microsystems Inc.  or
Java Version: 1.3.1 from Apple Computer, Inc.  or
Java Version: 1.1.5 from Netscape Communications Corporation
```

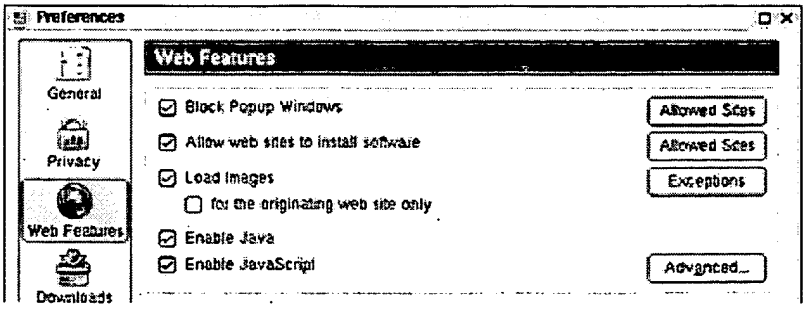
For actual results, see expected output.

Example of enabling Java in Internet Explorer (Windows)

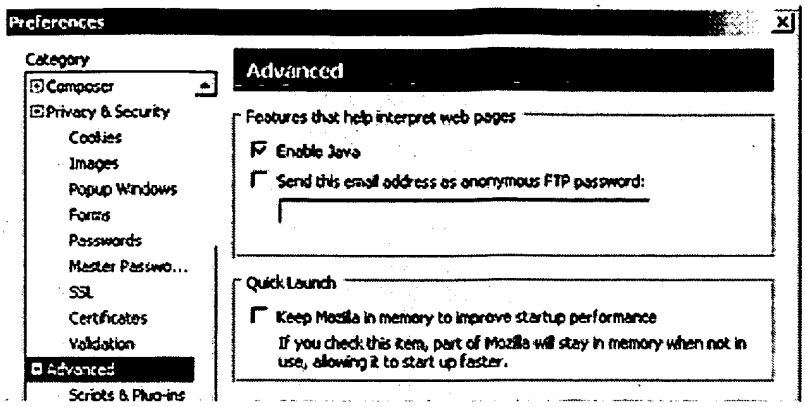


Examples of Enabling Java in Firefox (Windows/Linux)

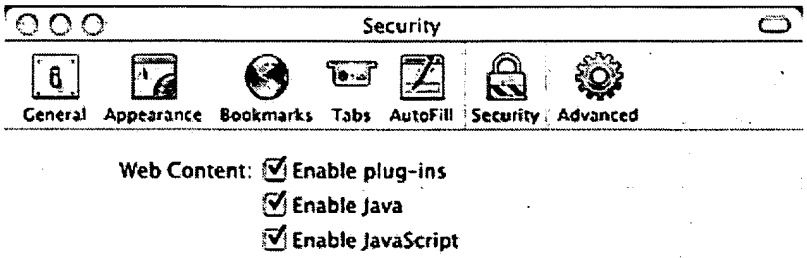




Example of Enabling java in Mozilla (Windows)

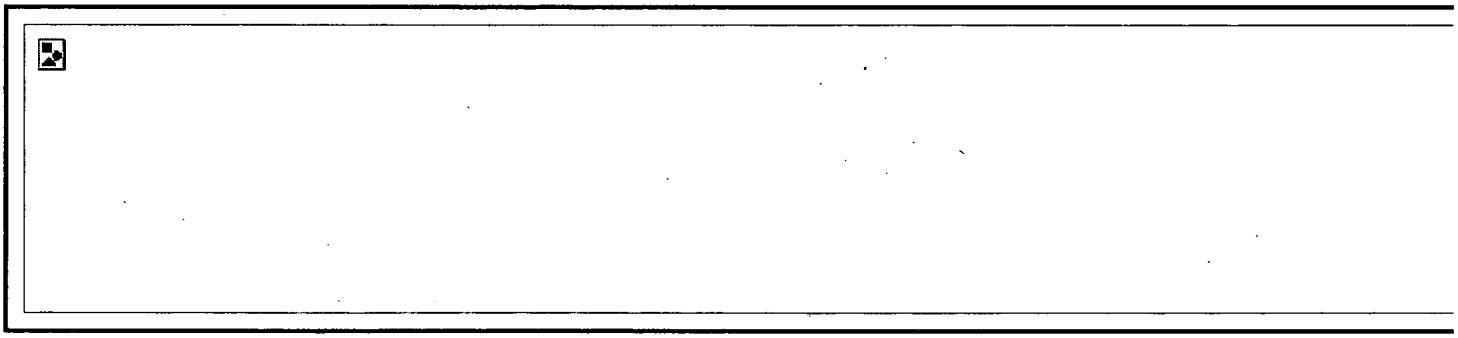


Example of Enabling java in Safari (Mac)



example:

Right-click in the area below for options!



IMPORTANT:
To use this option you will need to input a sequence alignment. Please make sure this alignment is in PIR or PHYLIP format. ALN and GCG MSF files are not supported so you will have to convert your MSF files to PIR format with, for example, GCG's ToPir:

```
topir pileup.msf{*} -outf=pileup.pir
```

Please refer to the GCG documentation to find out how to use this program correctly. You may then use this file as input (cut and paste or upload) to this service.

The method used is the NJ (Neighbour Joining) method of Saitou and Nei. First you calculate distances (percent divergence) between all pairs of sequence from a multiple alignment; second you apply the NJ method to the distance matrix.

This option allows you to choose the following output formats for the tree:

- Neighbour
- Phylip
- Distance

In order to view these trees you must have a program capable of displaying the data. Please refer to this pages section on OUTPUT for more information.

- **Kimura Correction of distances**
  This options allows you to set on distances correction (correction for multiple substitutions). This is because, as sequences diverge, more than one substitution will happen at many sites. However, you only see one difference when you look at the present day sequences. Therefore, this option has the effect of stretching branch lengths in trees (especially long branches). The corrections used here (for DNA or proteins) are both due to Motoo Kimura.

- **Ignore Gaps in alignment**
  With this option, any alignment positions where ANY of the sequences have a gap will be ignored. This means that 'like' will be compared to 'like' in all distances. It also, automatically throws away the most ambiguous parts of the alignment, which are concentrated around gaps (usually). The disadvantage is that you may throw away much of the data if there are many gaps.

• UPLOAD A FILE

You may upload a file from your computer which containing a valid **set** of sequences in any format (GCG, FASTA, EMBL, GenBank, PIR, NBRF, Phylip or UniProtKB/Swiss-Prot) using this option. Please note that this option only works with Netscape Browsers or Internet Explorer version 5 or later. Some word processors may yield unpredictable results as hidden/control characters may be present in the files. It is best to save files with the Unix format option to avoid hidden windows characters. Some examples of common sequence formats may be seen here.

• SCORES TABLE

Scores Table is a new view to ClustalW output. Users can sort the scores by Alignment Score, Sequence Number, Sequence Name and Sequence Length.

• REFERENCES

Higgins D., Thompson J., Gibson T. Thompson J. D., Higgins D. G., Gibson T. J. (1994).
CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting,position-specific gap penalties and weight matrix choice.
Nucleic Acids Res. 22:4673-4680. VIEW

ClustalW WWW Service at the European Bioinformatics Institute
http://www.ebi.ac.uk/clustalw
Rodrigo Lopez, Services Program

Andrew Lloyd
The Clustal W WW server at the EBI
embnet.news volume 4.2 1997
http://www.ebi.ac.uk/embnet.news/vol4_3/clustalw1.html

The program in use in this service can be obtained freely from:
http://www.sgi.com/chembio/resources/clustalw/parallel_clustalw.html

Parry-Smith, D.J., Payne, A.WR, Michie, A.D. and Attwood, T. K. (1997)
CINEMA - A novel Colour INteractive Editor for Multiple Alignments.
Gene, 211(2), GC45-56. VIEW
http://www.bioinf.man.ac.uk/dbbrowser/CINEMA2.1/

Attwood, T. K., Payne, A. W.R., Michie, A.D. and Parry-Smith, D.J. (1997)
A Colour INteractive Editor for Multiple Alignments - CINEMA,
EMBnet.news, 3 (3).

Jalview - a java multiple alignment editor
http://www.jalview.org

pfaat (Protein Family Alignment Annotation Tool)

http://pfaat.sourceforge.net/

- **OTHER SERVICES:**

This services is also available as an application from the EBI's srs server:
http://srs.ebi.ac.uk/

N.B. DbClustal can be launched from Wu-Blast2 and NCBI-Blast2 you can choose some of the sequences from your score list and align them with your query sequence. You may also then paste your .aln files back into ClustalW (bootstrapping) in order to colour your alignments, view the phylogenetic trees or launch jalview.